



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 130752

TO: Barba Kroma
Location: REM-2B692C18
Art Unit: 1638
Friday, September 03, 2004

Case Serial Number: 10/732721

From: Deirdre Arnold
Location: Biotech-Chem Library
REM 1A64
Phone: 571-272-2532

Deirdre.Arnold@uspto.gov

Search Notes

Thank you for using STIC services.

This Page Blank (uspto)

Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches. Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions **.nrpm** and **.nrpn**. Searches run against the Amino Acid Pending database produce two sets of results, with the extensions **.rapm** and **.rapn**.

Because they contain data that is confidential, the results of Pending database searches should not be left in the case.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 06:48:28 ; Search time 6301 Seconds
(without alignments)
11404.975 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattcagcggtcaatac.....ctcgttagtttggaacggcg 1658

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.man.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

*ed. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	68.8	4.1	2000	6	AX655393	AX655393 Sequence
c	64.6	3.9	150918	2	BX784041	BX784041 Danio rer
2	60.6	3.7	200337	9	AL590608	AL590608 Human DNA
3	59.8	3.6	78101	8	AF031569	AF031569 Zea mays
4	59.8	3.6	346296	8	AF090447	AF090447 Zea mays
5	56.8	3.4	78756	2	PFMAL6P1_13	Continuation (14 o
6	56.4	3.4	273275	3	AE014828	AE014828 Plasmodiu
c	56.2	3.4	1791	3	AF151389	AF151389 Dermatobi
7	55.8	3.4	1792	3	AF151388	AF151388 Dermatobi
8	55.8	3.4	328	11	BV074228	BV074228 S212P6542
9	55.8	3.4	125316	9	AC092635	AC092635 Homo sapi
10	55.8	3.4	335050	3	PFA929356	AL929356 Plasmodiu
11	55.8	3.3	205429	2	AC005506	AC005506 Plasmodiu
12	55.8	3.3	253132	3	AE014846	AE014846 Plasmodiu
13	55.8	3.3	313050	3	PFA929352	AL929352 Plasmodiu
14	54.6	3.3	83902	5	AL672161	AL672161 Zebrafish
15	54.4	3.3	200337	9	AL590608	AL590608 Human DNA
16	54.4	3.3	302156	3	AC116977	AC116977 Dictyoste
17	54.4	3.3	111882	3	AC115612	AC115612 Dictyoste
18	54.2	3.3	313050	3	PFA929352	AL929352 Plasmodiu
19	54.2	3.2	2750	6	A43608	A43608 Sequence 33
20	53.6	3.2	12779	3	AE001381	AE001381 Plasmodiu
21	53.6	3.2	110000	2	PFMAL7P1_04	Continuation (5 of
22	53.6	3.2	110000	2	PFMAL7P1_05	Continuation (6 of
23	53.6	3.2	114475	9	AC108082	AC108082 Homo sapi
24	53.6	3.2	196248	2	AC020725	AC020725 Homo sapi
25	53.6	3.2	251065	2	AC115145	AC115145 Rattus no
26	53.4	3.2	110000	2	PFMAL8P1_06	Continuation (7 of
27	53.4	3.2	250029	3	AE014820	AE014820 Plasmodiu
28	53.4	3.2	250029	3	AE014839	AE014839 Plasmodiu
29	53.4	3.2	7498	6	AX345159	AX345159 Sequence
30	53.2	3.2	176212	2	AC141778	AC141778 Apis mell
31	53.2	3.2	192929	2	AC005505	AC005505 Plasmodiu
32	53.2	3.2	250713	3	AE014850	AE014850 Plasmodiu
33	53.2	3.2	58697	2	AC144975	AC144975 Silurana
34	53.2	3.2	141752	2	BX537131	BX537131 Danio rer
35	53.2	3.2	234081	3	PFMAL4P2	AL035475 Plasmodiu
36	53.2	3.2	103343	9	AC108078	AC108078 Homo sapi
37	52.8	3.2	110000	2	PFMAL6P1_04	Continuation (5 of
38	52.8	3.2	111882	3	AC115612	AC115612 Dictyoste
39	52.8	3.2	114276	9	AC011355	AC011355 Homo sapi
40	52.8	3.2	250029	3	AE014816	AE014816 Plasmodiu
41	52.6	3.2	172497	2	BX276093	BX276093 Danio rer
42	52.6	3.2	193113	5	AL935197	AL935197 Zebrafish

ALIGNMENTS

RESULT 1	AX655393	Sequence 5263 from Patent WO03000898.	2000 bp	DNA	linear	PAT 22-MAR-2003
LOCUS	AX655393					
DEFINITION	AX655393					
ACCESSION	AX655393					
VERSION	AX655393.1	GI:29158207				
KEYWORDS						
SOURCE						
ORGANISM						
	Oryza sativa					
	Oryza sativa					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;					
	Ehrhartoideae; Oryzaceae; Oryza.					
REFERENCE						
AUTHORS						
	Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,					
	Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.					

```

TITLE      Plant genes involved in defense against pathogens
JOURNAL    Patent: WO 03000898-A 5263 03-JAN-2003;
            Syngenta Participations AG (CH)
FEATURES   Location/Qualifiers
            source
            1..2000
            /organism="Oryza sativa"
            /mol_type="unassigned DNA"
            /db_xref="taxon:4530"
ORIGIN
Query Match      4.1%; Score 68.8; DB 6; Length 2000;
Best Local Similarity 11.8%; Pred. No. 9e-05;
Matches 81; Conservative 310; Mismatches 292; Indels 3; Gaps 2;
Qy 397 AGATGTACAGTGCAGCAAGCTTCGATTTCACATGACGAGTAAATTTTAAATGACAGGA 456
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 362 MRWEMWTRRRRWAQKSSRSRKRKMKWKRYKRMGRYSRMRSCKRARWKRCHSG 421
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 457 GGCCTTTGATTCCTCAGCAGCTCTCGCGCCCGTCTGCTGTGACGCACTGACGAGG 516
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 422 RAWMGRCGCMTCRMKSYGMRWMSKRMASKYKWMRMYRWKXKCSRTTWWGTRGGM 481
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 517 CAGCAGCAGCTCTAGTGCACAGTCTGCTGCTGCTGCCACCGTATGATGCGAAGGTTG 576
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 482 MGTMGRCYKRSKMRKRCRRRRWRMYRMEWKRYMSARYTMYRCARKKYSYSAARKAR 541
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 577 CTTCCTTGCCTGCTGCTGCTGCGAGCAGCAGGTTGAAGTACGCGGCTGCTTGC 636
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 542 CWYRGKYWAGWMMKRYKMYMYMMWYKRYKSKSWYCKMYASCMKSAFKAGAKM 601
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 637 CTGCGCGCTGTAGCAGCGGACGACGCTGACGCTGATGCTGCTGCTGCTGCTGCTG 696
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 602 CKRSKMSAWSMSSSRCKRCKASRSKRYAMMGWTSRMSRWSKSYCYWRKWSGM 661
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 697 TTGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 754
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 662 KSTCTMYYMSKYFYAKYGSYRWYRYRAWCMYMWYRYRYRYRYRYRYRYRYRYRY 721
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 755 TGATCGACGTGTGCTGCGGCGATTTCAAAATCAGTAAATCAGTGTGATCCCATGACT 814
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 722 GMKSGRYWTSWYKCYKSKYRSMWYVWVWVWVWVWVWVWVWVWVWVWVWVWVWV 781
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 815 TCGTGTGCGGCAATAATCTGCTAGGATTTGCTTTTCAATGATGCTGCTGCTGCTG 874
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 782 WCTMWGYWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWV 841
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 875 GGAAGAATAATATACATAAATA-TGAGTTCAACATCTTCTTAATACCTTTTGTCTG 933
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 842 RGRARKARYWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWV 901
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 934 ATCTCCTTAGTAGAGCCAGCTGCTGATAACGCTGTTAAGAACCCCTGTTTACCGAGTGA 993
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 902 RYKTCWKAWSWYVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWV 961
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 994 GTCCAGGCTGCTTAATGATTAATAGTAAACACAGGAGACAGAGTGTAGTAGGAGAA 1053
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 962 CCAGWAWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWVWV 1021
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1054 CTGATCTCTTGTACTATATGTTGCT 1079
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1022 CCTACACACTACACAGAAATGTTGT 1047
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
BX784041/c
LOCUS      BX784041.1
DEFINITION Danio rerio clone DKEYP-95D10, WORKING DRAFT SEQUENCE, 12 unordered
            pieces.
ACCESSION  BX784041
VERSION     BX784041.1 GI:38304074
KEYWORDS   HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE     Danio rerio (zebrafish)
ORGANISM   Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
Burton, J.
Direct Submission
Submitted (10-NOV-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zkp95D10
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 145795 bases at least Q40
Consensus quality: 147142 bases at least Q30
Consensus quality: 148218 bases at least Q20
Insert size: 149818; sum-of-contigs
Insert size: 169833; 6.1% error; agarose-fp
Quality coverage: 4.15x in Q20 bases; agarose-fp
Quality coverage: 3.89x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 2197: contig of 2196 bp in length
* 2196: gap of 100 bp
* 2297: contig of 24511 bp in length
* 26807: gap of 100 bp
* 26808: contig of 9421 bp in length
* 36328: gap of 100 bp
* 36329: contig of 18035 bp in length
* 54463: gap of 100 bp
* 54464: contig of 10560 bp in length
* 65123: gap of 100 bp
* 65224: contig of 7188 bp in length
* 72411: gap of 100 bp
* 72512: contig of 2584 bp in length
* 75096: gap of 100 bp
* 75196: contig of 15120 bp in length
* 90316: gap of 100 bp
* 90415: contig of 5469 bp in length
* 95884: gap of 100 bp
* 95885: contig of 27986 bp in length
* 123970: gap of 100 bp
* 124071: contig of 3379 bp in length
* 127450: gap of 100 bp
* 127550: contig of 23369 bp in length.
*
* Location/Qualifiers
* 1..150918
* /organism="Danio rerio"
* /mol_type="genomic DNA"
* /db_xref="taxon:7955"
* /clone="DKEYP-95D10"
* /clone_lib="Daniokeypilot"
* 1..2196
* /note="assembly fragment:00049
* fragment_chain:1"
* 2297..26807
* /note="assembly fragment:00741
* fragment_chain:1"
* 26908..36328
* /note="assembly fragment:00183
* fragment_chain:1"

```


Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 78101)

Song,R., Llaca,V., Linton,E. and Messing,J.

Sequence, regulation, and evolution of the maize 22-kD alpha zein

gene family

Genome Res. 11 (11), 1817-1825 (2001)

21549079

MEDLINE

2 (bases 1 to 78101)

Llaca,V. and Messing,J.

Direct Submission

Submitted (27-OCT-1997) Waksman Institute, Rutgers, The State

University of New Jersey, P.O. Box 759, Piscataway, NJ 08855-0759,

USA

FEATURES

source

Location/Qualifiers

1..78101

/organism="Zea mays"

/mol_type="genomic DNA"

/db_xref="taxon:4577"

/chromosome="4"

/map="4s"

/note="cosmid contig; inbred line BSSS53"

1..39083

/organism="Zea mays"

/mol_type="genomic DNA"

/db_xref="taxon:4577"

/chromosome="4"

/map="4s"

/clone="III.4H10"

/note="inbred line BSSS53"

25530..68272

/organism="Zea mays"

/mol_type="genomic DNA"

/db_xref="taxon:4577"

/chromosome="4"

/map="4s"

/clone="III.5D8"

/note="inbred line BSSS53"

40872..73554

/organism="Zea mays"

/mol_type="genomic DNA"

/db_xref="taxon:4577"

/chromosome="4"

/map="4s"

/clone="V.3B1"

/note="inbred line BSSS53"

41990..78101

/organism="Zea mays"

/mol_type="genomic DNA"

/db_xref="taxon:4577"

/chromosome="4"

/map="4s"

/clone="V.9C11"

/note="inbred line BSSS53"

1..541

/note="similar to Zea mays c1 locus myb homolog CDS and

gpc4 promoter region"

1471..2270

/gene="azs22-1"

/note="22-kDa alpha zein 1"

/pseudo

5582..6382

/gene="azs22-2"

/note="22-kDa alpha zein 2"

/pseudo

9689..9766

/gene="azs22-3"

/note="22-kDa alpha zein 3; truncated"

/pseudo

10702..10770

/note="microsatellite"

/rpt_type=tandem

repeat_region

repeat_region

/rpt_unit="AT"

15639..15694

/note="microsatellite"

/rpt_type=tandem

/rpt_unit="AT"

18119..18919

/gene="azs22-4"

18119..18919

/gene="azs22-4"

/codon_start=1

/product="22-kDa alpha zein 4"

/protein_id="AAC01573.1"

/db_xref="GI:2832243"

/translation="MATKILSLALLALFASATNAFIPOCSLAPSSIIITOLFPEVTS

MGFHPAVQAYRLQQAIAASVLOQPISQLOQSLAHLTIOTIATCQQOQFIPLASHLA

MVNPAYLQQQLASNPALANVAVNGQQQLQQLFALPSALMWNPAAYLQQOQLLS

SSPLAVANAPYLOQLLQIVPALTOLVANPAAYLQQLLPFLNLTMSNAAYLQQR

QQLNPLAVANPLVAAFLOQQLLYNQSLINPVLRSQPIVGAIF"

23325..24125

/gene="azs22-5"

/note="22-kDa alpha zein 5"

/pseudo

27044..28675

/rpt_family="PREM-1 retroelement-related sequence"

/rpt_type=dispersed

30446..31840

/codon_start=1

/product="hypothetical protein"

/protein_id="AAC01574.1"

/db_xref="GI:2832244"

/translation="MMVQIVVHRAKTSTDFLVKKEAVSSTPQVNPVPVSGTAAVHAAR

QHTAOTSRTSPRAAGALLAARELLRHPSSSTASPGAKQKQDDVNRQDMAHSTST

RSRQSSHRYEASTSMRSPVGRQNTDLVNLNHRAGADAVSLERACRRONIE

GRNLDDPAEVVPPQAPMGTRSQGVPLAGVCAALVDHLRAASWSPKRPPLPEKIDG

TSNPSEFLQVYVITAAAGNTAMATYFHALSGFARTLWLNLSPGSIYSWEELIAR

FVANFASAYQQHGVGAEHLHVAHQEPRETRLMFISFTKVGQTIPTISDASITAPRQG

VRDEKMLEKLATHDVEIVPTLLALADKCAAAEGRAHMLAPQTGATGSGGSAIPRDG

KKKKKDRDHOKPRSTALVAVGTGGRGNRPRPQGRNGSGSCPVHNGRHTVECR

ELDLAKRVSRHE"

36867..37673

/gene="azs22-6"

/note="22-kDa alpha zein 6"

/pseudo

40977..42480

/gene="azs22-7"

/note="22-kDa alpha zein 7"

/pseudo

41499..42213

/note="insertion"

misc_feature

complement(48352..51063)

/note="similar to retrovirus-related POL polypeptide

sequence"

complement(49176..49850)

/codon_start=1

/product="hypothetical protein"

/protein_id="AAC01575.1"

/db_xref="GI:2832245"

/translation="MQYLATIRSLERQPKFTLQHVDRKNEEADALAKAARGAALP

SDVFYHVTGTPAFRSPGELQINDSEGHRIVNLIMTDMRAPITLFLQSYIYPTDINE

AKCLKHSRDFALIEQYIKGVQPMKVCVTETEGVQLIREVHSGTCGSHAGPRALA

AKVIHQGFYWPAMICAAANRVTRSCACQKFSRSGSPSQFTKLIATHTWFLQRMGLDVI

GELPTA"

52998..53798

/gene="azs22-8"

52998..53798

/gene="azs22-8"

/codon_start=1

/product="22-kDa alpha zein 8"

/protein_id="AAC01576.1"

/db_xref="GI:2832246"

/translation="MATKILALLALLFVSATNAFIIPQCSLAPSAIIPQLRPVTS

MGFHLAVQAYRLQQAASVLOQPIINQLOQSLAHLTIQTATCQQOQFLPALSQLD

VVNPAAYLQQQVLASNPALANVAAYQQQQLQQLFALPSALMWNPAAYLQQOQLLS

SSPLVGNAPTYLQOQLLOQIVPALTOAVANPAAYLQQLLPFNQLTVNSAGYLQOR
QQLNPLAVNELVTAFLQOQLLPYSQFSIMNPALSWQQPVGGAIF"
58815..59610
/gene="azs22-9"
/note="22-kDa alpha zein 9"
/pseudo
complement(67340..68552)
/note="similar to Zea mays globulin-1 promoter"
misc_feature
68037..68563
/note="similar to Zea mays globulin-1 promoter"
misc_feature
71597..72372
/note="similar to Zea mays zrp2 promoter"
gene
75093..75893
/gene="azs22-10"
CDS
75093..75893
/gene="azs22-10"
/codon_start=1
/product="22-kDa alpha zein 10"
/protein_id="AAC01577.1"
/db_xref="GI:2832247"
/translation:"MATKILALLALLFLVSATNNAFIPOCSLAPSAIIPOFLPPVTS
MGHEHAVQNRQQLAAVLSVQPIAQLOOQSLAHITIOAITTQOQQOFLPALSHIA
MVNPAAYLQOQLLALANVANVQOQLOOFLPALSQLAMVNPAAVYVQOQQLIS
SSPLVGNAPTYLQOQLLOQIVPALTOAVANPAAYLQQLLPFNQLTLSNSAAYLQOR
QQLNPLVIANPLVAFLQOQLLPYNQFSIMNPVLARQQPVGGAIF"
ORIGIN
Query Match 3.6%; Score 59.8; DB 8; Length 78101;
Best Local Similarity 62.3%; Pred. No. 0.027;
Matches 160; Conservative 0; Mismatches 92; Indels 5; Gaps 4;
QY 942 AGTAGACCCACGTCGTGATAACGTGTTAAGAACCCCTGTACCGAGTGTAGTCCAAGG 1001
Db 60320 AATGGACATTCGTGCTGATAACGTGTTAAGAACCCCTGTAGTGTAGTCCAAGG 60379
QY 1002 CTGCTTAACCT- GAATAATAGTAGAACACAGGACAGAGGTGTAGAGGAACTGATT 1059
Db 60380 CTTCTTAATAGAAATAGAAATCATAACACAGGACACAAATGTATAGGAGAACTGTT 60439
QY 1060 CTTTGTACTATATGTCGTCTCTCAAGGTTACATGATATGGGATCTCTCTCTAT 1119
Db 60440 ATTCTTATCATGTATT-CAGCTCANTGGTTACAAACAAATGGTGTCTCTCTCTATAT 60498
QY 1120 TTATAGACAAACTAGGTTTTCAGGCATATGGCCACATAGGCCTTCTCGGCCCAAGAAA 1179
Db 60499 ATAGGCT-AGAATAGAGTTTCGGCTAGATGGCCACATGGCCCTTCGGGGCCCAA-ACT 60556
QY 1180 GGTTCCTTAACACTACC 1196
Db 60557 GGTTCCTTAACACTCCC 60573
RESULT 5
AF090447
LOCUS Zea mays 346296 bp DNA linear PLN 14-MAR-2003
DEFINITION Zea mays 22 kDa alpha zein gene cluster, complete sequence.
ACCESSION AF090447 AF105716
VERSION AF090447.2 GI:13606087
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 346296)
REFERENCE
AUTHORS Song,R., Liaca,V., Linton,E. and Messing,J.
TITLE Sequence, regulation, and evolution of the maize 22-kD alpha zein
gene family
JOURNAL Genome Res. 11 (11), 1817-1825 (2001)
MEDLINE 21549079
PUBMED 11691845
REFERENCE
2 (bases 80292 to 116863)
AUTHORS Song,R., Liaca,V. and Messing,J.

Direct Submission
Submitted (04-SEP-1998) Waksman Institute, Rutgers University, 190
Frelinghuysen Road, Piscataway, NJ 08854-8020, USA
3 (bases 1 to 65155)
Liaca,V., Lou,A., Young,S. and Messing,J.
Direct Submission
Submitted (10-NOV-1998) Waksman Institute, Rutgers University, 190
Frelinghuysen Road, Piscataway, NJ 08854-8020, USA
4 (bases 1 to 346296)
Song,R., Liaca,V. and Messing,J.
Direct Submission
Submitted (12-APR-2001) Waksman Institute, Rutgers University, 190
Frelinghuysen Road, Piscataway, NJ 08854-8020, USA
Sequence update by submitter
5 (bases 1 to 346296)
Song,R., Liaca,V. and Messing,J.
Direct Submission
Submitted (24-APR-2001) Waksman Institute, Rutgers University, 190
Frelinghuysen Road, Piscataway, NJ 08854-8020, USA
Amino acid sequence updated by submitter
On or before Apr 12, 2001 this sequence version replaced
gi:4416300, gi:4140643.
FEATURES
source
1..346296
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
/chromosome="4"
/map="4s"
/clone="cosmid II.2E10"
/notes="inbred line BSS53"
1..65155
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
/chromosome="4"
/map="4s"
/clone="cosmid III.3C12; V.9D7"
/notes="inbred line BSS53"
1261..185036
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
/chromosome="4"
/map="4s"
/clone="BAC 204"
/notes="inbred line BSS53"
80292..116863
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
/chromosome="4"
/map="4s"
/clone="cosmid II.2E10"
/notes="inbred line BSS53"
172910..346296
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
/chromosome="4"
/map="4s"
/clone="BAC 171"
/notes="inbred line BSS53"
join(<1256..1397,2128..>2300)
/product="hypothetical protein"
join(1256..1397,2128..2300)
/codon_start=1
/product="hypothetical protein"
/protein_id="AAD20309.1"
/db_xref="GI:4416304"
/translation="MWKLLGVLEHDDTDADRSSEKGLDVGFTEETTEQESTGVRVYK
KAGSSATVPSSAQKVTVMRMEFFSLKCLLPHSIKDQNTGWTFRVNDGTELIRLQ
IR"

```

mRNA
join(<3179..4286,5020..5155,5266..5443,5892..>6020)
/product="hypothetical protein"
CDS
join(3179..4286,5020..5155,5266..5443,5892..6020)
/codon_start=1
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="AAD20310.1"
/db_xref="GI:4416305"
/translation="MVQKGTVEGSDTSVLEQTQTVGASVSNLAQCCTCSAVFGDD
KEMANCKVEHSGASRTVIVASGNGHTETACGGCGSNCNPMMAEDPPKAGNAKSGG
CSGCGCGGAGCGCTWANEGQSRSGVGTLLSNAMGEGQQAHSAGCGSGCGGG
SGMWIEGSKANNAMSGCGCGCGGGLI EYSKANNAMSGCGCGCGGCGSGM
VNEGSKANNAMSGCGCGCGGCGGALFNASTAAGELNKNKSAAGCGCGCGGGGG
CGSNALIGFKANAKSGCGCGCGGCGGCTLFSSAAAGQSGRSGCGCGCGGCGS
GSMVAEGNGHAKSGCGCGCGGCGGCGGATTLTSLMPGADVTAQESGTA
RCEELVPAFTPTFTFYGGKASLQRFACGDCGDELSPFLMPVETMEYIHN
YIAPDSRLHARDIYNPOHDFVETCCCSSTSPWLAASLQNAFEGTKBQTAATQ
TCCCCFQK"
join(<1112..11154,11262..11396,12254..12344,14272..14410,
14838..14890,18185..18230,18589..>18657)
/product="hypothetical protein"
CDS
join(1112..11154,11262..11396,12254..12344,14272..14410,
14838..14890,18185..18230,18589..18657)
/codon_start=1
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="AAD20311.1"
/db_xref="GI:4416306"
/translation="MDVGMWPMATRVDRVERSAHGDGLTGMVGATGEAETHGGG
BEVEGQHDMAQREVLRDMIGRSTRTCSVGTWSVAPPRGLARYITEPLLEMSF
DHRPCYAGLGVETPEKRTVLDPSPHYCESATNLYEALDNRDGLSRQLLRSI
GMEDPKVTLVKNQBAISLMSIKDPQAAK"
15142..17012
/note="similar to cell division protein FTS2"
22068..22142
/rpt_type=dispersed
24303..24307
/note="sequence target duplication"
24306..31758
/rpt_family="zeon-1 retroelement"
24308..24976
/rpt_type=dispersed
<25532..>26569
/product="gag protein"
25532..26569
/codon_start=1
/product="gag protein"
/protein_id="AAD20308.1"
/db_xref="GI:4416303"
/translation="MSYQATVSSYGNTAVMAKSFVMVRSVAQTWYSSLRPGTITSW
QKLDMLTSGFGFTQKPTAQLFCQTDHEYLQAVRRFLRLRAQPTVPNEIVI
EAMIKGLRPGSAQYFARKPPTLKLQKMDXEYIRADNFRREERAFRSEMTRGF
GGFYFVRVIRHNSNDKSGSQOQRPQCSQASQOQSFPFPAPRGRGARGFER
FGDQPRIRICLFGKNGKHTRMCHVTIQKQKEIAEAAQAQPKQIMHTASYHSPYI
PEYGNHPAVSVASASQASWHQPPPPPLQQQOQPEGQVAQHQDRFREQSEARTV
NSTVPESKHIV"
31090..31758
31759..31763
/note="sequence target duplication"
32800..32855
/rpt_type=dispersed
36025..37602
/note="pip(F10)20725 RFLP marker"
join(<36233..36579,37382..>37925)
/product="hypothetical protein"
join(36233..36579,37382..37925)
/codon_start=1
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="AAD20312.1"
/db_xref="GI:4416307"
/translation="WGSFKGHVLPGLFLAVGLWRVWSAARHAAEPSSFRVWSPA
PAPGPARULELVVAGGALADMCVEVLYSTHLRIIPAGGEVNPALHNDLEHGMLLMEF

```

```

LFGALASQLWPRHPLTDGALCLVAATAFTAEVLVFFHSTHMGLEGYHYLLVL
LVGCVAAALVIGALLPDSFPADLASGLVLAQGLWIFYOTLALALYGMILPAGCARDADA
HVECSRRAAQRAEQALASQLFALVLLAFYALGCVAAVAAAYGRPELPAVRHRAA
AVELECRQDDAGAMECAI"
complement(38797..38910)
/rpt_type=dispersed
complement(39147..39151)
/note="target sequence duplication"
complement(39152..49223)
/rpt_family="gypsy/Ty3-type retrotransposon"
/rpt_type=dispersed
complement(39152..42406)
complement(39152..45283)
/note="gypsy-type polyprotein"
/pseudo
/codon_start=1
complement(45973..49223)
complement(49224..49228)
/note="target sequence duplication"
complement(50243..50571)
/note="similar to grande 1"
complement(50572..60252)
/note="nested in grande 1 retroelement"
/rpt_family="copia-type retroelement PREM-2"
/rpt_type=dispersed
complement(50572..50576)
/note="target sequence duplication"
complement(50577..52000)
/note="copia-type retroelement PREM-2"
complement(<52949..>56140)
/product="copia-type pol polyprotein"
/codon_start=1
/product="copia-type pol polyprotein"
/protein_id="AAD20307.1"
Query Match 3.6%; Score 59.8; DB 8; Length 346296;
Best Local Similarity 62.3%; Pred.No.0.039; 92; Indels 5; Gaps 4;
Matches 160; Conservative 0; Mismatches 92; Indels 5; Gaps 4;
QY 942 AGTAGAGCCACGTGGTGATACGTTAAGAACCCCTTTGTTACCGAGTGTAGTCCAAGG 1001
DB 206658 AATGAGCATTCGTGCTGATACGTTAAGAACCCCTTTGTTACCGAGTGTAGTCCAAGG 206717
QY 1002 CTGCTAACT--GAATAATAGTAGAACACACAGACAGAGTGTAGAGAGAACTGATT 1059
DB 206718 CTTCTTAATTAGAAATAGAAATCATAACACAGAGACACAAATGTATAGGAGAACTGTT 206777
QY 1060 CTTTCTTACTATATGGTGTCTCTCCAAAGGTTACATGATATGGGATCTCTCTCTAT 1119
DB 206778 ATTCTTATTCATGTATT-CAGCTCAATGTTTACAACAATGGTGTGTCTCTCTATAT 206836
QY 1120 TTATAGACAAACTAGGGTTTCAGGCATATATGGCCACATAGGCCCTTCTTGGCCCCAAGAA 1179
DB 206837 ATAGGAT-AGAATTAGGGTTTCGGTAGATGGCCACATAGGCCCTTCTTGGCCCCA-ACT 206894
QY 1180 GGTTCCTTAACACTACC 1196
DB 206895 GGTTCCTTAACACTCCC 206911

```

RESULT 6 PFMAL6P1_13 WPCOMMENT

Sequence split into 14 fragments LOCUS PFMAL6P1 Accession AL844505

Fragment Name	Begin	End
PFMAL6P1_00	1	110000
PFMAL6P1_01	100001	210000
PFMAL6P1_02	200001	310000
PFMAL6P1_03	300001	410000
PFMAL6P1_04	400001	510000
PFMAL6P1_05	500001	610000
PFMAL6P1_06	600001	710000


```
repeat_region /note="AT rich"
               /rpt_type=tandem
               complement(2591..2614)
repeat_region /rpt_type=tandem
               /rpt_type=tandem
               /rpt_unit="(A)n"
               complement(2625..2656)
repeat_region /note="AT rich"
               /rpt_type=tandem
               complement(2662..2690)
repeat_region /rpt_type=tandem
               /rpt_type=tandem
               /rpt_unit="(A)n"
               complement(2778..2807)
repeat_region /note="AT rich"
               /rpt_type=tandem
               complement(2808..2856)
repeat_region /rpt_type=tandem
               /rpt_type=tandem
               /rpt_unit="(TA)n"
               complement(2808..2852)
repeat_region /rpt_type=tandem
               /rpt_type=tandem
               complement(2870..2929)
repeat_region /rpt_type=tandem
               /rpt_type=tandem
               /rpt_unit="(TAAAA)n"
               complement(2966..2987)
repeat_region /rpt_type=tandem
               /rpt_type=tandem
               /rpt_unit="(A)n"
               complement(3049..3070)
repeat_region /note="AT rich"
               /rpt_type=tandem
               complement(3078..3150)
repeat_region /rpt_type=tandem
               /rpt_type=tandem
               complement(3078..3142)
repeat_region /rpt_type=tandem
               /rpt_unit="(TA)n"
               complement(3206..3454)
repeat_region /note="AT rich"
               /rpt_type=tandem
               complement(3297..3376)
repeat_region /rpt_type=tandem
               /rpt_type=tandem
               complement(3297..3376)
repeat_region /rpt_type=tandem
               /rpt_type=tandem
               /rpt_unit="(TA)n"
               complement(3481..3504)
repeat_region /note="AT rich"
               /rpt_type=tandem
               complement(3570..3597)
repeat_region /rpt_type=tandem
               /rpt_type=tandem
               /rpt_unit="(A)n"
               complement(3654..3692)
repeat_region /note="AT rich"
               /rpt_type=tandem
               complement(3811..3859)
repeat_region /rpt_type=tandem
               /rpt_type=tandem
               /rpt_unit="(GAAAA)n"
               complement(3868..3928)
repeat_region /note="AT rich"
               /rpt_type=tandem
               complement(3949..3977)
repeat_region /rpt_type=tandem
               /rpt_type=tandem
               complement(4011..4056)
repeat_region /note="AT rich"
               /rpt_type=tandem
               complement(4187..4288)
repeat_region /rpt_type=tandem
               /rpt_type=tandem
               complement(4187..4288)
repeat_region /rpt_type=tandem
               /rpt_type=tandem
               /rpt_unit="(TA)n"
               complement(4187..4282)
repeat_region /rpt_type=tandem
```

```
repeat_region complement(4290..4339)
               /note="AT rich"
               /rpt_type=tandem
               complement(4356..4391)
repeat_region /note="AT rich"
               /rpt_type=tandem
               complement(4395..4436)
repeat_region /rpt_type=tandem
               /rpt_type=tandem
               complement(4395..4436)
               /rpt_type=tandem
               /rpt_unit="(TA)n"
               <4567..>5065
gene          /locus_tag="PF14_0709"
mrna          join(<4567..4713,4847..>5065)
               /locus_tag="PF14_0709"
CDS           join(4567..4713,4847..5065)
               /locus_tag="PF14_0709"

Query Match      3.4%; Score 56.4; DB 3; Length 273275;
Best Local Similarity 51.6%; Pred. No. 0.22;
Matches 129; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 65 TAATCATTTTCAGAGGATGAATAAATAATCGCAAGAAAGCAAAATATTTTAAATGAATGATG 124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 163990 TATTATTGCAATAACGGGAAATACCAAAAAAATAAATTCAAACAAITTTAA 163931

QY 125 CAATATACAAATTTTAAATTACACAATTATGTAAGATTACATTTGTTAGTTTCATAGAAATC 184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 163930 AAATATATATATATATATATATTTTATTATTATACGTTTATATAAATATATAT 163871

QY 185 AATTTCTAGAGTCATAATAATGCCTAAATAAATCTATCTATTTTCCCTTAAGAAA 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 163870 AAAATATATTTACCTATATATTTTAAAAAATAAATTTCTATATCATTTTAAATTA 163811

QY 245 AAAACAGCCATTAAAGGACCATTAGATGCGCGCTGCCATTTTATATATATATATATAT 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 163810 TATGCTCAACCTTATGATATATATAAATAAGCAATATATATATATATATATATAT 163751

QY 305 TATGAGTTGT 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 163750 AATGAGTTAT 163741
```

RESULT 8

AF151389/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

AF151389 1791 bp DNA linear INV 30-SEP-2003
Dermatobia hominis crRNA-Ile Gene, partial sequence; and 12S
ribosomal RNA, partial sequence; mitochondrial genes for
mitochondrial products.
AF151389
AF151389.1 GI:7159047
mitochondrion Dermatobia hominis (human botfly)
Dermatobia hominis
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Oestroidea; Oestridae; Dermatobia.
1 (sites)
Lessinger, A.C. and Azeredo-Espin, A.M.
Evolution and structural organisation of mitochondrial DNA control
region of myiasis-causing flies
Med. Vet. Entomol. 14 (1), 71-80 (2000)
20221065
10759315
2 (bases 1 to 1791)
Lessinger, A.C. and Azeredo-Espin, A.M.L.
Direct Submission
Submitted (14-MAY-1999) Centro de Biologia Molecular e Engenharia
Genetica - CBMEG, Universidade Estadual de Campinas - UNICAMP, Lab.
Genetica Animal, CBMEG, UNICAMP, P. O. Box 6109, Campinas, Sao
Paulo 13083-970, Brazil
Location/Qualifiers
1..1791

D-loop	/product="tRNA-Ile"
tRNA	43..1612
	1613..>1792
ORIGIN	/product="12S ribosomal RNA"
Query Match	3.4%; Score 56; DB 3; Length 1792;
Best Local Similarity	49.3%; Pred.No. 0.073;
Matches 146; Conservative	0; Mismatches 150; Indels 0; Gaps 0;
QY	12 TCACAATACAGTCATCTACATGTGATAAATTATATAGTTAGATTACTTCTTGTAATCAT 71
DB	1249 TCAATATTTATATATCATATATTAATATTTAAATATTTAAATTAATTAATAATCAC 1190
QY	72 TTCAGAGGATGAAGAAAAAATCGCAAGAAGCAAAATATTTTAAATGAATGATGCAATATA 131
DB	1189 TAATAAAAACCTACATATAAATCAAACAATAAATTTATAAATTTTAAATTAATTTATAAAT 1130
QY	132 CAAATTTTAAATCACACAATATATGTAAGATACATGTTGTTTAGTTTCATAGAAATCAATTTCT 191
DB	1129 TATTATATTTATATAAATAAATAATATAAATAATATAATAATTTAAATTAATTAATACTTT 1070
QY	192 AGAGTCATATAATATGCCTAAATTTAAATTTCTATTTCTATTTTTCCCTTAAGAAAAAACAG 251
DB	1069 TTAATTTTAATAFAAAATAAATTTATAATAATAAATTTAAAGTAATATATGTATATACATAT 1010
QY	252 CCATTAAGGGACCATTAGATCGCGGCTGCCATTTTATTATATATAGAGATAT 307
DB	1009 ATAAATATATATTTGATTTATATATATATATAATTTAAATAATTTATAAATATATATATAT 954
RESULT 10	BV074228/c
LOCUS	BV074228 328 bp DNA linear STS 31-MAY-2003
DEFINITION	S212P6542FG4.T0 CZECHII/Bi Mus musculus STS genomic, sequence tagged site.
ACCESSION	BV074228
VERSION	BV074228.1 GI:31190023
KEYWORDS	STS.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Theria; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 328)
AUTHORS	Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C., Lander,E.S., Lindblad-Toh,K. and Daly,M.J.
TITLE	The mosaic structure of variation in the laboratory mouse genome
JOURNAL	Nature 420 (6915), 574-578 (2002)
MEDLINE	22354684
PUBMED	12466852
COMMENT	Contact: Kerstin Lindblad-Toh Whitehead Institute for Biomedical Research, Center for Genome Research 320 Charles Street, Cambridge, MA 02141, USA Tel: 6172521477 Fax: 6172580903 Email: kersli@genome.wi.mit.edu Primer A: None Primer B: None STS size: 328 Protocol: WGS-discovery: Paired-end low-coverage whole genome shotgun reads were generated from 129S1/SvimJ, C3H/HeJ, and BALB/cByJ. The WGS reads were placed uniquely on the MGSCv3 C57BL/6J assembly and SNP detection was carried out by SSAHA-SNP. 225,000 reads were annotated as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J and the strain from which the particular read came. The validation rate for these SNPs was estimated at approximately 98%. Location/Qualifiers 1..328 /organism="Mus musculus"
FEATURES	source

source	Location/Qualifiers
1. .125316	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone="RP11-342E23"
	/clone_lib="RP11-11"
43. .137	/rpt_family="Achobo"
397. .636	/rpt_family="L1"
800. .841	/rpt_family="AT_rich"
1476. .1623	/rpt_family="MIR"
1943. .2100	/rpt_family="MIR"
2207. .2310	/rpt_family="L2"
2356. .2404	/rpt_family="Mariner"
2447. .2496	/rpt_family="Mariner"
2506. .2649	/rpt_family="L1"
2916. .2976	/rpt_family="ERV1"
2964. .3035	

```

repeat_region /rpt_family="(TA)n"
3058. .3162 /rpt_family="(TA)n"
repeat_region /rpt_family="(TA)n"
3207. .3230 /rpt_family="(TA)n"
repeat_region /rpt_family="AT_rich"
3215. .4207 /rpt_family="AT_rich"
repeat_region /rpt_family="L1"
4220. .4504 /rpt_family="L1"
repeat_region /rpt_family="Alu"
4626. .4917 /rpt_family="Alu"
repeat_region /rpt_family="ERV1"
5260. .5425 /rpt_family="ERV1"
repeat_region /rpt_family="L1"
5661. .5998 /rpt_family="L1"
repeat_region /rpt_family="ERV1"
6019. .6191 /rpt_family="ERV1"
repeat_region /rpt_family="(TGA)n"
6226. .6264 /rpt_family="(TGA)n"
misc_feature 6666. 6802 /note="match to EST AA603518 (NID:G2437379) np07dl2.s1"
repeat_region 8592. .9336 /rpt_family="L2"
repeat_region 9882. .9909 /rpt_family="L2"
repeat_region 10148. 10169 /rpt_family="AT_rich"
repeat_region 10365. 10397 /rpt_family="AT_rich"
repeat_region 11309. 11355 /rpt_family="AT_rich"
repeat_region 11354. 11504 /rpt_family="AT_rich"
repeat_region 11500. 11576 /rpt_family="L1"
repeat_region 11552. 11735 /rpt_family="L1"
repeat_region 11865. 12101 /rpt_family="L1"
repeat_region 11988. 12025 /rpt_family="L1"
repeat_region 12148. 12286 /rpt_family="AT_rich"
repeat_region 12354. 12476 /rpt_family="L1"
repeat_region 12453. 12504 /rpt_family="L1"
repeat_region 12477. 12784 /rpt_family="T-rich"
repeat_region 12785. 15212 /rpt_family="Alu"
repeat_region 14498. 14532 /rpt_family="L1"
repeat_region 15706. 15920 /rpt_family="AT_rich"
repeat_region 15925. 17405 /rpt_family="L1"
repeat_region 16747. 16770 /rpt_family="L1"
repeat_region 17406. 17774 /rpt_family="AT_rich"
repeat_region 17775. 18148 /rpt_family="MaLR"
repeat_region 18234. 18257 /rpt_family="L1"
repeat_region 18518. 18557 /rpt_family="AT_rich"
repeat_region 18900. 18932 /rpt_family="(CA)n"
repeat_region 18993. 19186 /rpt_family="(TG)n"
repeat_region 20216. 20677 /rpt_family="MER1_type"

```

```

repeat_region 20996. 21268 /rpt_family="L1"
repeat_region 21307. 21979 /rpt_family="L1"
repeat_region 21590. 21633 /rpt_family="L1"
repeat_region 21988. 22117 /rpt_family="AT_rich"
repeat_region 22172. 22794 /rpt_family="L1"
repeat_region 22638. 22659 /rpt_family="L1"
repeat_region 22638. 22659 /rpt_family="AT_rich"

```

```

Query Match 3.4%; Score 55.8; DB 9; Length 125316;
Best Local Similarity 52.3%; Pred. No. 0.25;
Matches 123; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 32 ATGTGATAAATTTTATAGTTAGATTACTTCTTGTAAATCAATTCAGAGAGTCAAAAAAAA 91
Db 94026 ATGCTAGGTTTGTATTACATACATACAGCATTAATTTGTTAGAAATAGCTTACAGGA 94085

QY 92 TCGCAGAAAGCAAAATTTTAAATGAATGATGCAATATACAAATTTAAATTAATCACAATTA 151
Db 94086 TAGAAAGAAATTTACTATATTACAAGATTCAACAATCTCGAAGAGTGTATAATATCA 94145

QY 152 TGTAAAGATTACATGTTTGTAGTTTCATAGAATCAATTTCTAGAGTATATAATGCTTAA 211
Db 94146 TTTGATAGTTCTTTATAGTAAGTTAAATACATTTATCTTTTAAATATATAAAATATTA 94205

QY 212 ATTAATAATCTATTCTATTCTTTTCTTAAAGAAAAAACACCCATTAAAGGACCA 266
Db 94206 ATAAAACTAATAATTAACITTAACATTCGAGATAGATGAATTCATTATAAAACCA 94260

```

```

RESULT 12
PFAS29356 335050 bp DNA linear INV 29-JAN-2003
LOCUS Plasmodium falciparum strain 3D7, chromosome 9; segment 2/5.
DEFINITION AL929356 AL844508
ACCESSION AL929356.1 GI:233504968
KEYWORDS
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 335050)
AUTHORS Hall.N., Pain.A., Berriman.M., Churcher.C., Harris.B., Harris.D., Mungall.K., Bowman.S., Atkin.R., Baker.S., Barron.A., Brooks.K., Buckee.C.O., Burrows.C., Cherevach.I., Chillingworth.C., Chillingworth.T., Christodoulou.Z., Clark.L., Clark.R., Corton.C., Cronin.A., Davies.R., Davis.P., Dear.P., Dearden.F., Doggett.J., Feltwell.T., Goble.A., Goodhead.I., Gwilliam.R., Hamlin.N., Hance.Z., Harper.D., Hauser.H., Hornsby.T., Holroyd.S., Horrocks.P., Humphray.S., Jagels.K., James.K.D., Johnson.D., Kerhornou.A., Knights.A., Konfortov.B., Kyes.S., Larke.N., Lawson.D., Lennard.N., Line.A., Maddison.M., McLean.J., Mooney.P., Moulé.S., Murphy.L., Oliver.K., Ormond.D., Price.C., Quail.M.A., Rabinowitz.E., Rajandream.M.A., Rutter.S., Rutherford.K.M., Sanders.M., Simmonds.M., Seeger.K., Sharp.S., Smith.R., Squares.R., Squares.S., Stevens.K., Taylor.K., Tivey.A., Unwin.L., Whitehead.S., Woodward.J., Sulston.J.E., Craig.A., Newbold.C. and Barrell.B.G.
TITLE Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13
JOURNAL Nature 419 (6906), 527-531 (2002)
MEDLINE 22255708
PUBMED 12368867
REFERENCE 2 (bases 1 to 335050)
AUTHORS Sanders.M., Hauser.H., Baker.S., Unwin.L., Mungall.K., Berriman.M., Pain.A., Hall.N., Bowman.S., Churcher.C., Quail.M. and Barrell.B.
TITLE Direct Submission
JOURNAL Submitted (20-SEP-2002) P.falciparum Genome Sequencing Consortium, The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
COMMENT For more information about this sequence or the Malaria Project,

```



```

CDS
/gene="PFI0355c"
/complement(21948..24716)
/gene="PFI0355c"
/notes="C-terminal half of protein is similar to
prokaryotic ATP-dependent heat shock proteins. Similar to
Rhizobium loti heat shock protein, ATP-dependent hsl
protease ATP-binding subunit, hslu ml15004 SMALL-Q98CU1
(EMBL:AP030005) (443 aa) fasta scores: E(): 8.5e-54,
46.69% id in 439 aa, and to Escherichia coli, and
ATP-dependent hsl protease ATP-binding subunit hslu or
htpi or b3931 or z5478 or ecs4858 SMALL:HSU ECOLI
(SWALL:P32168) (443 aa) fasta scores: E(): 2.1e-51, 46.56%
id in 451 aa. No HMPfam hit; SMART hit to SM00382,
ATPases associated with a variety of cellular activities"
/codon_start=1
/product="ATP-dependent heat shock protein, putative"
/protein_id="CAD51757.1"

Query Match 3.4%; Score 55.8; DB 3; Length 335050;
Best Local Similarity 55.4%; Pred. No. 0.32; Indels 1; Gaps 1;
Matches 128; Conservative 0; Mismatches 102;

QY 30 ACATGTGATAAAATTTTAGTGTAGTACTTCTTGTAATCAATTTTCAGAGATGCAAAAAA 89
Db 168695 AATTTTAATTAATGATAGCATATATATAAATATAACAATCAAAAAA 168754

QY 90 AATCGCAAGAAAGCAATATTTTAATGAATGATGCAATATACAAATTTTATACACAAT 149
Db 168755 AGAAATTAATTAAGGGAAAAATATATAGGAATTTATTTATATATATATATATAT 168814

QY 150 TATCTAAGAT-TACATTTGTTTAGTTTCATAGAAATCAATTTCTAGAGTCATAATAATGCC 208
Db 168815 AATTATAATATATATATATATATATATATATATTTTATTTTATTTTATTTATTTATTT 168874

QY 209 TAAATTTAAATPTCTATCTATTTTTCCTTAAGAAAAAAGAGCCCAATTA 259
Db 168875 TATATTAAATTTTTTTTTTTTTTTTTTTTTTTTAAATTTTATAAGTAGTAA 168925

RESULT 13
AC005506 205429 bp DNA linear HTG 12-AUG-2000
LOCUS Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
DEFINITION PROGRESS ***, 4 unordered pieces.
AC005506
VERSION AC005506.8 GI:9797717
KEYWORDS HTG; HTGS PHASE1.
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM Plasmodium falciparum
REFERENCE Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
1 (bases 1 to 205429)
AUTHORS Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tamaki,T.,
Kurdi,O.B., Conway,A.B. and Davis,R.W.
Plasmodium falciparum 3D7 chromosome 12
Unpublished
2 (bases 1 to 205429)
AUTHORS Hyman,R.W., Fung,E.L., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.
Direct Submission
TITLE Submitted (21-AUG-1998) Stanford DNA Sequencing and Technology
JOURNAL Center, Stanford University, 855 California Avenue, Palo Alto, CA
94304, USA
COMMENT On Aug 12, 2000 this sequence version replaced gi:8810449.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 132269: contig of 132269 bp in length
* 132270 132469: gap of unknown length
* 132470 192227: contig of 59758 bp in length

```

```

* 192228 192427: gap of unknown length
* 192428 203864: contig of 11437 bp in length
* 203865 204064: gap of unknown length
* 204065 205429: contig of 1365 bp in length.
FEATURES
Location/Qualifiers
source 1..205429
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/db_xref="taxon:5833"
/chromosome="12"
/clone="PFYAC357"
/clone="3D7"
ORIGIN
Query Match 3.3%; Score 55; DB 2; Length 205429;
Best Local Similarity 48.8%; Pred. No. 0.43;
Matches 148; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 7 ACGGCTCACAAATACCAAGTCATCTACATGTGATAAAATTTTAGTGTAGTACTTCTTGTA 66
Db 79666 ACTGTAATGAATCAATATATTTTATTTATACCTATTGCTATATTATTACTAC 79725

QY 67 ATCATTTACAGAGATGAAAAAATCGCAAGAAAGCAAAATTTTAAATGAATGATGCA 126
Db 79726 GTCAAGTGAATATATGTAATAGAGAAATATATATATATATATATATTTTAAATGAC 79785

QY 127 ATATACAAATTTTATACACAATTTCTAAGATTCATCTTTAGTTTCATAGAAATCAA 186
Db 79786 ATATGAATAAATAATTTTANGTATATAAACAATATAATATATATATATATATATA 79845

QY 187 TTCTAGAGTCATAAATCGCTAAATTTAAATTTCTATTTTCTTTTAAAGAAAAA 246
Db 79846 TATATATATATATATATATATATATATATATATATATTTTATTTTATATATAAT 79905

QY 247 AACAGCCCATTAAGGGACCAATAGAAATGCGCGTCTCCATTTTATATATATAGAGATA 306
Db 79906 TATAGAAAAATAAATAGAAAGTAAACACAGTACAGTTTCAGTATTTGAAAGGAACACTA 79965

QY 307 TGA 309
Db 79966 TGA 79968

RESULT 14
AE014846/c
LOCUS Plasmodium falciparum 3D7 chromosome 12, section 3 of 9 of the
DEFINITION complete sequence.
AC005506 AE014846 AE014188
VERSION AE014846.1 GI:23496645
KEYWORDS
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
REFERENCE Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
1 (bases 1 to 253132)
AUTHORS Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W.,
Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T.,
James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A.,
Kyes,S., Chan,M.S., Nene,V., Shallom,S.J., Sub,B., Peterson,J.,
Angiuoli,S., Portea,M., Allen,J., Selengut,J., Haft,D.,
Mather,M.W., Vaidya,A.B., Martin,D.M., Fairlamb,A.H.,
Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I.,
Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C.,
Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and
Barrell,B.
Genome sequence of the human malaria parasite Plasmodium falciparum
JOURNAL Nature 419 (6906), 498-511 (2002)
MEDLINE 2255705
PUBMED 12368864
REFERENCE 2 (bases 1 to 253132)
AUTHORS Hyman,R.W., Fung,E., Conway,A., Kurdi,O., Mao,J., Miranda,M.,
Nakao,B., Rowley,D., Tamaki,T., Wang,F. and Davis,R.W.
Direct Submission
TITLE

```

JOURNAL	Submitted (13-SEP-2002) Stanford Genome Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA	CDS	complement (5840..7003) /locus_tag="PFL0570c"
REFERENCE	3 (bases 1 to 25132)		/codon_start=1
AUTHORS	Hyman,R.W., Fung,B., Conway,A., Kurdi,O., Mao,J., Miranda,M., Nakao,B., Rowley,D., Tamaki,T., Wang,F. and Davis,R.W.		/product="hypothetical protein"
TITLE	Direct Submision		/protein_id="AAN36203.1"
JOURNAL	Submitted (29-JAN-2003) Stanford Genome Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA		/db_xref="GI:23496647"
FEATURES	Location/Qualifiers		/translation="MIVLLKNNLKNLTLLRPLGCGNNKLVFMRGVINEQTKGPIEKKT TDNNEKMLNKTLEEDVKVSKGNCNALFDECEELYRMKKVKDKESVYDNPSPILBEL NREELKITRNIFERINKKKKHVSNSLYKDIYIDPYNPFKVDNLFKVEYDNPSPILBEL NASIVELKKQRAIKKSLKEQKLYNPYSEENKHNFLNMDNDVKEQMECHDKQEKY WNPSFKKILNLIKSPFIWRYTHLHPHIGENGLILPKINMTTKQOIQIFKSCICIA RMALYPYDKKPEKDDLPLMDPLOFLDELHRYIQLNDLPAQALIKVMINRPSLN YKVFYFSKKEKDKNEKTAQNEQENVESEKPPSQMLQYKKNYENFYK"
source	1..25132 /organism="Plasmodium falciparum 3D7" /mol_type="genomic DNA" /isolate="3D7" /db_xref="taxon:36329" /chromosome="12"	repeat_region	7326..7363
repeat_region	136..182	/rpt_type=tandem	
repeat_region	328..838	/rpt_type=tandem	7629..7664
repeat_region	1400..1425	/rpt_type=tandem	7668..7702
repeat_region	1556..1583	/rpt_type=tandem	8021..8048
repeat_region	1994..2022	/rpt_type=tandem	8092..8133
repeat_region	2249..2282	/rpt_type=tandem	8274..8311
repeat_region	2553..2591	/rpt_type=tandem	8355..>17469
repeat_region	2885..2909	/rpt_type=tandem	/locus_tag="PFL0575w"
repeat_region	2970..3091	/rpt_type=tandem	join(<8355..17090,17320..>17469)
repeat_region	3136..3162	/rpt_type=tandem	join(8355..17090,17320..17469)
gene	<3376..>4694		/locus_tag="PFL0575w"
mRNA	join(<3376..3481,3717..3818,3990..4253,4432..>4694)		/note="HMPfam hit to PF01593, Flavin containing amine oxidase, score 1e-05"
CDS	join(3376..3481,3717..3818,3990..4253,4432..4694)		/codon_start=1
repeat_region	3500..3544		/product="hypothetical protein"
repeat_region	3873..3905		/protein_id="AAN36204.1"
repeat_region	3917..3956		/db_xref="GI:23496648"
repeat_region	4817..4887		/translation="MGTEKEDQDIVNRNDLKKKKNNKCDNMSNDISOSISNIS DSISTNSNNISNNISNNISNNISNNISNNISNNISNNISNNISNNISNNISNNIS ISQNHDIIRKVVHILKNLIMNTHKNNYKNGNLINYYKKKQYGDENKILNTPRES YDFSYLKLSEKSLKQSDMIKONMSQDFINVFQKLLKHHNLDLIMNNKLYIYI KFSODIDLIPYFKERPFKTLIGNNILLAKKYEVTEVLKSLYKLTQKPTSTN FTIDELVHNNPMKYELIYRIVINPFPIOMNKNKGHKKKINNNNKKKYIDHE DINNVEYHDLBEENQKISSLPFKCYKNIKCMNFEKSHNNMKIOTVEDFOKENT ENGKILYLKEKGINNYNEKDMCMHSAFNNMVQNNVLDNMMCNKNTKKKKKN NVIFPMVDNYINDVSNKNHNSLOENNYSVLQIPNNKQINNIRANVILNNEF NLDVQKFRNTEETWNKIESKNDEERLQFEKKEKQILKREKLRNQEKKRQKQEK ENKQEKIYIKKELLILKEIRKILKEKELKEKEKKEKKEKFLKNEKEIKKIKED KKRQKEPLRKIKSEKKLKEKKLKEELINSIKEERKPKELKSKPKKSKKDNHL NKDKKTKQNKIDKSNQNNKKYNNNDKKDDINYLQFLIDQINNNTVDKKKKKTH DISLSIKKKKCTKNLLNSINDIKSDLSFISFTPTFNNIEKEQEDRNDINDILKL IEIKVYNTQEKNDYIECLHKSQKSNVRNKDKIQNQLLKGNNKIITKNGEHNDQ SDPLNQLNDLDDFLNNMTIIRKKYKGGTSDVMSVNDVDSNVDISNVDPSNKG ISNKGISVDISKNVHTNNDNIDNHYNNKNNNSERNESIEINELILKGLKSPG DINKLSLOKILIRIKESGKNKIEYVHI PNNVIERSAKKILVNLKEGIIKQKQVKE KYTKKISLPGIVKPEDIIRKGRNVNTHNEDNQLYDVLIIIGAGISGLAASLYLN KCNKFLVIEGRNIGGRFAFTILPERIINNKLVPETVDIGANYLHCCDADLTNDK KGTGKHFANCTEFNNEKSKEDDIYEDNFIWMKKKTKKKYKIVKRRKTYN KNGNKNKMGNFSTIDVQKEYVTNINLFEYKSLINDFAYLPFYVNNLSQSEHN QDDENKTDKNKISGGKNIVRYLNSKNKISCSNINYNORVSDDEDYFYSRRIKNRR KIKLRNNMLPMIKNNIDIRNYFKTYNKEKDNILNFPKSLDELIHFVNNLSDCVHT SSVLYPKNCTNKSNNKLLILSPKOTRRRREYDKSVTQLAELKPLSVLYSGKDNWE STFYAHYNNNGKKIKSKFYIRVNLCDKILRVAARMRNFIFIGSCNDEYQKIK SEDOQKNNYVSDTTEKDSYDYNNDILKEDQIINNLIIDDTYHKLIFNNHVK NIKPNENNIIRNENFEASVLYNNKDDMLENEENGKNAENGKENGENGKNAENGK EONEQKNGKNGEENGKNGEENGKNGKNGKNGKNNKTSKKHIISSKKKMS KYDANNILSYDNHAIINYNYDYDGLNVKVVNVSTNRNIDDERKKKEDISKIT ELYNNKKRRSMDDMECTEIEFTENNEENFTMBEWMMLMTVLSRYSGYGDLE TSIAMSLLPFSNLDIDMCPKYGNNYILKNIKYDYIKKNEQTPYISPFKDLNYSK IYVDGKWLINLYSEVQNKIFLNTVAELVYIKDHQDKKIYLDLNDNQINKQKSSS PLNTKKNNNNNNNNNNIESVNNNISNSHVINODNNSFVPONEVSSKNKKYINEF LINNVIKDDYKLGNEEDDYKVMIKCKYDTFKKINKHGHGTSDLNLCNKNYTIYA KYVIVALPLGCLIENTDKKRLKTCLEPELHPLKRALQNYKMGNNKNIILIRFPYN
repeat_region	3500..3544		/codon_start=1
repeat_region	3873..3905		/product="heat shock protein DnaJ homologue Pfj4"
repeat_region	3917..3956		/protein_id="AAN36202.1"
repeat_region	4817..4887		/db_xref="GI:23496646"
repeat_region	4820..4874		/translation="MSRRVNYEVLGPQADLTIVIKSKYRILAMKHPDKPNNAE ATERFKQISEAYELVSDPKRRKYDLYTDENYADENDEFNHFKNFGPNDQAQRIE MFGDSSPFGNDSFSDVMGSSFVDKRRGRVPRSDPNFPGSSFNVSFGSSPDNEF DGGSCFTSVETSTSGGKFKNRVVKTSKSTSIINGKRVTRIVTKLTPNGTVERTV TERBEDDRGNINIRQLPAHELRRNR"
repeat_region	5108..5144		/rpt_type=tandem
repeat_region	5319..5381		/rpt_type=tandem
repeat_region	5565..5636		/rpt_type=tandem
gene	complement(<5840..>7003)		/locus_tag="PFL0570c"
mRNA	complement(<5840..>7003)		/locus_tag="PFL0570c"

```

FWPDSLOINCIDQKQFQFLNLHAYKIGICILVHCPPMWSCTGYIKKEHYIINECLY
TLNKMERGKGLPILVYIITKWQDDNFSFSGIATYPIYINCNDNDLIYRSEPHINNP
RVVFCGEYLSKSFQCVGDYDTGIRAAEDIAHIGLHKNDTKNTKYTDVYIFPPQNK
PTNIPLPPIKFKLGFYITDGSDEALTDESSDENMNTVDINDINSNTNIDINN
NTNSSTSSNNNTKYNVNNPIPISWKKEVEFLNLSQSIQNIFFLKNQWTHKFN
KODKNIHFKETDALKCSYKKEIKSNTVESFTNLSLNFEDKSDNVKSNISTLL
NLYEYNNKTDQIHFDRIKSDSYKSKYKNGSYLLELSKSESDNSNDANNICTI
EFTNVDDKNDGHNVDNDVNDKYSNTNIEYIYSHITNNHFEYINNSISLY
INKYMTKUNDAPFYNTEIENILNKQNAEYKSYFVLLQTFNIFPKYSIDRFTYIQ
KELFKIQAMKHILFENQFQSDNMVLSQAENKNDHVSHNNHNDVDRKAYEQ
NTDIINEDDKETISYPHEKMEVEVNNMKLSLNLKSNMVGNNNDNDNNNGHIOKS
NLSTDELDYVILHLSKNTLTKNDKVMIOIYNNIDLONKNIYNNKINRKC
VLANDLYKYLILHLSKNTLDBOILSSNNKWNKSCGFLFELCVLOYILKHLKYDLFN
SVNTKVYQLCLDIYIYILFKHDLICYKNGGELNMCDSYNNCNSWHYCYLSSE
HNNNNKDFWLPSCNSIDISKHVQRCYQNEINENYWKRIYIYIKTFPUSRTK
IKRLDLKHLKHSRVA"
repeat_region 8427..8510
/rpt_type=tandem
repeat_region 8428..8498
/rpt_type=tandem
repeat_region 8510..8652
/rpt_type=tandem
repeat_region 10256..10314
/rpt_type=tandem
repeat_region 11019..11081
/rpt_type=tandem

Query Match 3.3%; Score 55; DB 3; Length 253132;
Best Local Similarity 48.8%; Pred. No. 0.45;
Matches 148; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 7 ACGGCTCACAATACCGATCATCTACATGTCATATAATTTATAGTAGTACTTCTGTGA 66
|||
Db 74935 ACTGTAATGAATCAATAATATTTATTTATTTATATACCTATTCCTATTTATTTACTAC 74876

QY 67 ATCATTTTCAGAGCATGAAAAAATCGCAAGAAAGCAAAATATTTTAAATGAATGATGCA 126
|||
Db 74875 GTGAGTGAATAATGTAATAGAGAAATATATATATATATATATATTTTAAATGAC 74816

QY 127 ATATACAAATTTAATACACAATATATGATAGATTACATGTTAGTTTCATAGAAATCAA 186
|||
Db 74815 ATATGAATAATAAATATTTAAGTATATAACATATAAATAATAATATATATATATATA 74756

QY 187 TTTCTAGATCAATAATATCGCTAAATTAATTAATCTATTTCTATTTTCTCTTAAAGAAAA 246
|||
Db 74755 TATATATATATATATATATATGATATATAATATTTTAAATTTTAAATATATATAAAT 74696

QY 247 AACAGCCCATTAAGGACCATTAAGATGCGCGTGTCTCCATTTTATATATATAGAGATA 306
|||
Db 74695 TATAGAAAAATAAATAGAGTAAGTAACAGTACAGTTTCAGTATTGAAAGGACACTA 74636

QY 307 TGA 309
|||
Db 74635 TGA 74633

```

```

RESULT 15
PFA929352/c
LOCUS Plasmodium falciparum strain 3D7, chromosome 5, segment 2/4.
DEFINITION AL929352 AL044504
ACCESSION AL929352.1 GI:23504570
VERSION
KEYWORDS Plasmodium falciparum 3D7
SOURCE Plasmodium falciparum 3D7
ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1
AUTHORS Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N.,

```

```

Hance, Z., Harper D., Hauser H., Hornsby T., Holroyd S.,
Horrocks P., Humphray S., Jagels K., James K.D., Johnson D.,
Kerhornou A., Knights A., Konfortov B., Kyes S., Larke N.,
Lawson D., Lennard N., Line A., Maddison M., McLean J., Mooney P.,
Moule S., Oliphant L., Oliver K., Ormond D., Price C., Quail M.A.,
Rabinowitsch E., Rajandream M.A., Rutter S., Rutherford K.M.,
Sanders M., Simmonds M., Seeger K., Sharp S., Smith R., Squares R.,
Squares S., Stevens K., Taylor K., Tivey A., Unwin L.,
Whitehead S., Woodward J., Sulston J.E., Craig A., Newbold C. and
Barrell B.G.
Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13
Nature 419 (6906), 527-531 (2002)
22255708
12368867
2 (bases 1 to 313050)
Devlin K., Baker S., Davies P., Mungall K., Berriman M., Pain A.,
Hall N., Bowman S., Churcher C., Quail M. and Barrell B.
Direct Submission
Submitted (20-SEP-2002) P.falciparum Genome Sequencing Consortium,
The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA, UK
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/Projects/P_falciparum.
FEATURES
Location/Qualifiers
1..313050
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosomes="5"
/complement(219..8039)
/gene="PFE0420c"
/complement(219..8039)
/gene="PFE0420c"
/notes="Similar to Plasmodium falciparum guanine
nucleotide exchange factor pfrccl (l327 aa) fasta scores:
E(): 9.6e-190, 97.459% id in l338 aa"
/codon_start=1
/product="guanine nucleotide exchange factor, putative"
/protein_id="CAD51450.1"
/db_xref="GI:23504571"
/db_xref="SPTREMBL:Q81418"
/translation="MYEQDEDDKDIHKYIFPPFKISTGLDHVACVAYVYKPKSIFTW
GNSCNQLGLGINRVFVNVTVVCFIVASICTNVATFVLKVNVDIDGCVSY
FGKNGNLGYKKRFLPHEHEKNTQKVEQMLQKVDPLNQLMDLAFGKKKE
KEKKKKKKNDIKKEKEENINNEKNNNNNNINSEKKKKNDINGHEAKYVQNK
KFFYNHFNKNDIGLPSDDIYDDYINISGESLLEADNISSEIDEDTNNEDVYNN
IYNDHTNEDQSDINETIDKSDNDYDDEKKKEKNIYKNEKLYSNITKENDKHDW
FTFPFKIVFPKNTKIKFTSCGDMHTLAICTGGVLYSWGYNNGFGTNGTQNVYEP
TPIFLPHSVDDNNDDDDDDNNNNNTYKPKKEDSIKDWKKNNVITCSGSKH
SLACLHGLDIYSGVGGNGRLGLGNIKSYNKPOLIKGLRNKRLIYVCAGTSHSGCIDS
NYYVTWNGKFKYLGHGNDIDIYYPKLVECLHNKKVKCMLSPFGFNSLALVKGDV
YWGTFNIYNNYVYISKLPKQINDIKYINIKYISHAISTYVCGITLVDGLISPGNKY
SYKEINDNDSDLLIKYLINKEVEEQHYGNDKSVHTKNNIISYIVKKDEHDI
NIPEKNKVLKSVKVIIDGSDHPSVFLGSKYVYTAGYNKNGELNGEFLKPKFGVP
ILLNICVNLKIACGVYVLCVSDVGLVYWGKNDKSGIYKIDFVPPVKNLT
NVTNIFAGDHSACIINKLLDKGTNGVNNVNVNGVNGVNNYSDECGDIYWGNAE
SGVKGIVGVDYTCIILLPKRINKIYKCSLGNHSLFTNTNEIFVCGSNNGNNG
LFFSKMVCNIVSVPTKVLHNNIYIKDLAGNTYSIISVDGFIYIFGFIKSGNG
SNMSCEVSLNIEYTNVYMGKYEHPFLYDNLKFGIGNNKQIILCDNKENNY
IKPKLITLYLTKENNKIIESIFSFHNVTFIOFNNSBIFAWGYTNNVHLGIGIPNIKY
LKHPKTIKTLWLYDQKLDYNSCSDMEINVTNNNSYNNIYINBITLVKNI
NTTELQNEYTYEGYEEIEFIYINVINNNIINCHLONLAKKEENNNMEYIKSF
EKDIIDISKHIEVLLNLIYEQKFNLYINQNFILISNANKEOHLNLLIPFSTY
IFQNRKLEQFIYIIQQQPVYLIILCLHNYKNNVKNQOLYKQKSYFYKNSNVLC
HNNINQHIKNNYNIINSYINHTYEQNDALKYNSKDIQNKYSFYKNSNVLC
FIIDLADPRNERVITFTFLIKLGEELKNSLNVSIFINIDTSIFFTLIQMLFFKK
EILINFAQCILDLDPNSFVTLDLQSRKVPKNIPLNQGVSNKMMQHTLQHTW
PMWGAHRENNLLINDMYNDVVDKNIHVKNSTHMDKNDGANKKKIKKKNNKKK
NNRKKNNNNNNNNNNNNNGGDDIYSGDDIYSGDDIYSGDDIYSGDDIYSGDDIY
EKDKEFLNFKGTQNESDKRFLVNDISCYKSGVFPNNIYININNEDENNYNNISQEI

```



```

ONENLNDNNVYVEINFCIPEELCKIFLNLKPPDMFKIILKELFKYFIIEYKLNDEI
NOTNKIYFYQNHHTIYIPFKLILMALINPILKMNENIAHFSYFNIFPHILNIRNKI
CDFLEILYKESLNLNINLIISIKYIEFHTFISLTVNINNIINTHHIYYKNDAYTNL
CFHEILNRSYVOLKLFQVCHIENLFRFQVLMLSFNDPALDVNFVYFYKNNIPIQ
KHKHHTLSTWGTWTKMDIOKERERHLNIGSGKDMGNIQPKGVSTFTGHODNKKI
NSMEGNEILLAKAKNDLNDRFPMKGMINGVMKRGDNQPMGCMGYNMTGNGPMGCVN
PMWGNPMGCMGYNMTGNGPMGCMGYNMTGNGPMGCMGYNMTGNGPMGCMGYNMTGNG
AFNLPTIKITSDNDKMKRKKBLIPSEDDIEFFICKLIYNNVLIRFLLEKKNWSI
CEFTKIPMQOYICRYKQTLIENNEYIFSIIRHKYENKNNIYIISCEKNCPILEHCTD
TNLLILKSLKMNWLSQDKENNLHLINKAIDIFLSDMETYVEFLFEPPANLSIY
KFKDDILOKSMQEIYFDMQNTYKDSFQKWRNIALQISLNLKKKGMNLYLK
LVEKQYIEDLIINIOHKKMOITVLKNAIIFVSKLYIEKPILLHSSYFKNNLFFVEL
KKKHHTQFKLPYNSSTVHIYOKDLINQIILSTHETLYTLVNVLTIEIFDLNNI
IKFSLILTKDNQNNIINEHTFTGSEIYMSVNSPYIYFPFKYSKTYLCSITGLHFMH
LJHNFVIDLY"
9222..10010
/gene="PFE0425w"
9222..10010
/gene="PFE0425w"
/note="Weakly similar to Arabidopsis thaliana F8n16.20
protein f8n16.20 TR:Q9ZV23 (EMBL:AC005727) (312 aa) fasta
scores: E(): 2e-12, 28.763% id in 299 aa"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD51451.1"
/db_xref="GI:23504572"
/db_xref="SPTREMBL:Q81417"
/translation="MPSKGRVMPADNRLPVSASLTKSEIWMKNSVGYDPYASYEIN
KKKKEDNEEDIKAKNLFNLSLRTGITSTTIFGACTVCHNIGHLPYQCRNFISLEKFD
INNMMNNEEDIKERKNLGLMSDDGSDSHSDSDSHSDSDSDSDSDSDSDSDSDS
HSDSNSHSDSYKKKHSHSHSHSKYKKKHSHSHSKNKKKHSHSHSKKHHSIENDTHK
SDEKHKYISSSKKKKIKKEKKKKYHSHKMKNYITVNSSD"
11261..15733
/gene="PFE0430w"
11261..15733
/gene="PFE0430w"
/note="Similar to Rattus norvegicus RNA helicase TR:Q62780
(EMBL:U25746) (1032 aa) fasta scores: E(): 2.4e-59,
38.387% id in 719 aa"
/codon_start=1
/product="ATP-dependent RNA helicase, putative"
/protein_id="CAD51452.1"
/db_xref="GI:23504573"
/db_xref="GOA:Q81416"
/db_xref="SPTREMBL:Q81415"
/translation="MINIYNIYKSKKEKKNDKDMDIKVKQDHDDSDVDLKRDN
ISSNKKKDVERTYIYNISYEEENKKKKKSKSYSDYSYNTKTKNIYNDHRTDE
YPSYAKKSSYHVHREKYPPDEKRYEHIKKNRYSYDYSNYDESDFSSTSY
YNERNKYYKYEDKKKKEKEKEKERKKRYHHTDDDDHLESYSSRSSRSIN
IKEKQKKLYLSHEKERKRYNSANSYSSYDETKERDKKKSYHHDVFLDDKKN
VHNYESDRETKDESNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
QEI CENKEIKQVENNSNVKVI TENNNLSRLERLU FARKIKTNDTITNEKHTKFT
LNTVKYKLNLI F INDQDESNDSEKFEFDIKENKINQVNDNNHKKDDTKKIE
NHKIKEDSNTNDIHWIENKNEEDALDFMKEIEQACEEEKYKVKSSITLDEIYTD
MKQYEPFNPKNNEEKIYNAKVGNNIQQNTHQDSNVNTQEDLKKENKNTK
NTNNYNSDNNNNDDDDDDDDVFKHFLVEELKKKTEEDMKENKNNESNLTD
SKNKQKNGSDASFSSDDTDDKLSNKSSELNENGMKKNNKLLLVNHDIDIFIPI
KKNIVQYKEITNMKDSVDFPKNGNLIIVRGKNCPRPVQFYQCGPLSKILQLEK
KNFKWMYNIQWGTIPALMGCDVIAIAETGSKTSLYFPIRVHLYHQEFLRNDGPI
SIILPTRELISQVNEAKIYCKAVNIILAVGSGNTARQLKVKGEIILVGPGR
IIDIILTSCKVTNRSFVFLDEADRLDLGFSQIYVNLIRNCRKDKQAMISATF
PNYIQNAKKLLYKPIELIIVGEGKTKNNIYQFVEIIEBSKVFLLLKLGEWIKGL
VLIFVNOKEADLLVLEYKYDYNLLVLLHGQDDQDTRFTLEKFKENKYNLIATVSM
ARGIDIKNILVINOCPDHIEDYIHRIGRTGSNNIYGAYTFILPNPETYKADLYNL
LKNNIYLYNKTLIDIPDLENMIIEYTKINSINEKQKGNKNLGYKGGYKFTPDEKSR
HQIDIALAKELQENDENEKQLQVDDELGNIENHNTNVCNLKADQNKYDQFE
DYKSNITYPYITHSNMMFEDKNGISKNMSKKNTEKKNTOYDITKVEQEIQR
IKNDTMELSKAKKILNMLKTYNLSLQENHDKFLKNAQYEEFDIAEKALKVTEH
IKDEKXKLIENKTKEDIKKLIQNKIQSDTNESIQRNMLLSALRTKMKKISPYLP
HYTITEDNNLIEEFFYNNIPQHRUKISHKOVAKIADMWGATCOIKGYSNPSQPNK
QNFLDITKQLHIEILAPTYNQIARIENLNLNNFMISCNIKAKRASTGWAPP"
complement (16588..17442)
/gene="PFE0435c"

```

CDS

```

complement (16588..17442)
/gene="PFE0435c"
/note="Similar to Rickettsia prowazekii single-strand
binding protein ssb or rp836 SW:SSB_RICPR (Q92CC2) (152
aa) fasta scores: E(): 1.1e-16, 38.158% id in 152 aa
Signal peptide predicted by SignalP 2.0 HMM (Signal
peptide probability 0.896, signal anchor probability 0.017)
with cleavage site probability 0.471 between residues 25
and 26"
/codon_start=1
/product="single-strand binding protein, putative"
/protein_id="CAD51453.1"
/db_xref="GI:23504574"
/db_xref="SPTREMBL:Q81415"
/translation="MGKRLMCIYFPLLIYFNVVLHRTYGIIGDVKXVHOLNNIINKRI
IKSRKISMFKINLQDNFENKRFYNNMNRNMVNEKSLNKLIMLIGRVGCEPDITKILG
GDKVATFSLATNEFMRDNRNTNELKSKTDWHRIVVYDQNTVDLIDKVLKRGRRVYVQGS
LHTRKWHITNDMSQPKQITTEILSYNKGDLIFLDDKRNQRRNNSNNINSENOQHINN
EHINNINNNGDNDFMLNSDKLIEDKEFTDLDDNNEENNFTQSNSETFDKQGIYDK
MNVQBFEE"
20022..29426
/gene="PFE0440w"
20022..29426
/gene="PFE0440w"
/note="Weak similarity to Plasmodium falciparum Pfc0425W
protein pfc0425W TR:O77336 (EMBL:Z98547) (4550 aa) fasta
scores: E(): 7.2e-23, 22.792% id in 3238 aa"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD51454.1"
/db_xref="GI:23504575"
/db_xref="SPTREMBL:Q81414"
/translation="MSTCNNIVDLVRNKKELLTGKMKRKKSCNIDRNTIDNENKLL
DKNLIINFLNSELKIKETQLNNISLRSATYLLKIEKIOKLISCNNWKOVPISYIK
KSAENILLKNEYIKNLTKLNKAQNINNDINIDNIYNSCVNLNKNDEKK
FSKHLNSSSSSLSSYSLNPSDDSCFNSKQNNPFNSIKNNSTFTGNAQTKNK
LDNHTSKTKKINDKIYDKIKDKKIDKIKYDKYDKYDKDKDKDKDKDKDKDKYD
KIYDKKDKYDKYDKIKDKDKDKDKDKDKDKDKDKDKDKDKDKDKDKDKDKDKDKDK
YSNKIINYKGFYDNNKNEIKNNEYIYDDQDLDSYNGLCVYNGKNNIFCLNPQE

```

gene

CDS

```

Query Match 3.3%; Score 55; DB 3; Length 313050;
Best Local Similarity 54.8%; Pred. No. 0.48;
Matches 109; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
QY 26 ATCTACATGTGATAAATTTTATAGTTAGTACTTCTTGTATCATTTTCAGAGGATGAA 85
DB 199656 ATATATATATATATATATATATTTTATTTTATTTTATTTTATTTTATTTTAT 199597
QY 86 AAAAAATCGAAGAAGCAATATTTTAAATGAATGATGCAATATACAAATTTAATTACA 145
DB 199596 AAAAAAAAAAAAAAAAAATTTTTTTTAAATTTTAAATCATTAATAAAAAAAAAA 199537
QY 146 CAATTATGTAGATTACATTGTTTAGTTTCATAGAAATCAATTTCTAGAGTCATATAAT 205
DB 199536 AAAAAATTAAGTTTATATCAATTTGTACATATTAATGTGAACATACAAAAAATGTAT 199477
QY 206 GCCTAAATTAATAATTCAT 224
DB 199476 GCATAAAAAAAAAAATAT 199458

```

Search completed: September 2, 2004, 09:38:04
Job time : 6308 secs

gene

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 06:07:12 ; Search time 639 Seconds

Title: US-10-732-721-1

Perfect score: 1658

Sequence: 1 gaattcacgggtcacatac.....ctcgtagtttgggacggcg 1658

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 5747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002s:*

7: Geneseqn2003as:*

8: Geneseqn2003bs:*

9: Geneseqn2003cs:*

10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	68.8	4.1	2000	7	ADA71938	Ada71938 Rice gene
2	53.2	3.2	7498	6	ABL32257	Ab132257 Human imm
3	52.2	3.1	6012	6	ABK31371	Abk31371 Signal tr
4	52.2	3.1	6012	6	ABL70328	Ab170328 Chemical
5	52.2	3.1	6012	6	AAS61275	Aas61275 Human gen
6	51.4	3.1	6265	2	AAX08523	Aax08523 NBP46 (ro
7	50.8	3.1	12763	6	ABL32303	Ab132303 Human imm
8	50.8	3.1	113515	6	ABL34174	Ab134174 Human imm
9	50.6	3.1	17280	4	AAS46772	Aas46772 Tumour su
10	50.4	3.0	1134	6	ABT07549	Abt07549 Human bre
11	50.4	3.0	15767	6	ABL33206	Ab133206 Human imm
12	50.4	3.0	15767	6	ABL33206	Ab133206 Human imm
13	50.2	3.0	8056	7	AB210246	Ab210246 Human met
14	50.2	3.0	17142	4	AAS45498	Aas45498 Haematopo
15	50.2	3.0	17142	6	ABL34107	Ab134107 Human imm
16	50.2	3.0	17142	6	ABK28430	Abk28430 DNA trans
17	50	3.0	8067	6	ABN80224	Abn80224 Human che
18	49.4	3.0	14551	6	ABL34584	Ab134584 Human met
19	49.2	3.0	1501	7	AB210042	Ab210042 Haematopo
20	49.2	3.0	1501	9	ADE84086	Ade84086 Human lym
21	49	3.0	2000	7	ADA71938	Ada71938 Rice gene
22	48.8	2.9	8961	6	ABK28428	Abk28428 DNA trans
23	48.8	2.9	8961	6	ABL49380	Ab149380 Human pol

C	24	48.6	2.9	5891	6	AAS61087	Aas61087 Human gen
	25	48.4	2.9	632	5	ABA06279	Soy bean
	26	48.4	2.9	736	5	ABA06275	Soy bean
	27	48.4	2.9	4590	1	AAN60472	Sequence
C	28	48.2	2.9	5313	6	ABL32763	Human imm
C	29	48.2	2.9	5768	6	ABL34163	Human imm
C	30	48.2	2.9	7276	6	ABL32902	Human imm
C	31	48.2	2.9	13038	6	ABL33275	Human imm
C	32	48	2.9	2418	2	AAQ27886	P.falcipa
C	33	48	2.9	6053	4	AAS45371	Chemical
C	34	48	2.9	6053	6	ABK28208	DNA trans
C	35	48	2.9	7657	4	AAS45477	Chemical
C	36	48	2.9	7657	6	ABL34022	Human imm
C	37	48	2.9	33053	6	ABQ67006	Human ang
C	38	47.6	2.9	5499	6	ABQ66972	Human ang
C	39	47.6	2.9	9963	6	ABL32694	Human imm
C	40	47.6	2.9	18997	6	ABL32570	Human imm
C	41	47.6	2.9	18997	6	ABK33948	Human DNA
C	42	47.6	2.9	18997	7	ADA20352	Prostate
C	43	47.6	2.9	18997	7	ADA84159	Human ren
C	44	47.6	2.9	113515	6	ABL34175	Human imm
C	45	47.4	2.9	6078	6	ABK28227	DNA trans

ALIGNMENTS

RESULT 1
ADA71938
ID ADA71938 standard; DNA; 2000 BP.
XX
AC ADA71938;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5263.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
XX gene expression.
PS Claim 27; SEQ ID NO 5263; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to

KW	CpG island; signal transduction associated disease; solid tumour; cancer;
KW	antitumour; cytostatic; mutant; ds.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
PN	W0200200926-A2.
XX	
PD	03-JAN-2002.
XX	
XX	29-JUN-2001; 2001WO-EP007472.
PF	
XX	
XX	30-JUN-2000; 2000DE-01032529.
PR	
PR	01-SEP-2000; 2000DE-01043826.
XX	
PA	(EPIG-) EPIGENOMICS AG.
XX	
PI	Olek A, Piepenbrock C, Berlin K;
XX	
DR	WPI; 2002-147896/19.
XX	
PT	Oligonucleotide for diagnosis and therapy of diseases associated with
PT	signal transduction e.g. cancer, comprises chemically modified genomic
PT	sequences of genes associated with signal transduction.
XX	
PS	Claim 1; SEQ ID NO 214; 24pp; English.
XX	
CC	The present invention relates to chemically modified DNA sequences of
CC	signal transduction associated genes. The DNA sequences are chemically
CC	modified using a solution of bisulphite, hydrogen sulphite or disulphite.
CC	Also disclosed are oligonucleotides and/or PNA oligomers for detecting
CC	the cytosine methylation state (CpG islands) of these genes, and a method
CC	for the diagnosis and/or therapy of genetic and epigenetic parameters of
CC	genes associated with signal transduction. The genomic DNA can be
CC	obtained from cells or cellular components which contain DNA, e.g. cell
CC	lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,
CC	tissue embedded in paraffin such as tissue from eyes, intestine, kidney,
CC	brain, heart, prostate, lung, breast or liver, histologic object slides,
CC	and all their possible combinations. The sequences of the invention are
CC	useful for the diagnosis and therapy of diseases associated with signal
CC	transduction e.g. solid tumours and cancer. ABK31158-ABK31545 represent
CC	chemically pretreated genomic DNA sequences of different genes associated
CC	with signal transduction, or their complementary sequences. Note: The
CC	sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from the
CC	European Patent Office
XX	
SQ	Sequence 6012 BP; 2120 A; 32 C; 1149 G; 2711 T; 0 U; 0 Other;
	Query Match 3.1%; Score 52.2; DB 6; Length 6012;
	Best Local Similarity 50.2%; Pred. NO. 0.023;
	Matches 129; Conservative 0; Mismatches 128; Indels 0; Gaps 0
QY	3 ATTCAACGGCTCACAAATACCACTATCTACATGTGATAAATTTTATAGTTAGATTACTTCT 62
Db	1299 ATTCAATTCGAATTCACCAAAATATACCAAAATTTATTATAAAATCTCTATTAT 124
QY	63 TGTAAATCATTTACAGAGGATGAAAAAATAATCGAAGAAGCAAAATATTTTAAATGAATGA 122
Db	1239 TCATTCCTCTTATCATTTCAACAAAAATAAACAACTATTAATTTTCAATA 118
QY	123 TGCAAATACAAAATTAATTACAAATATCTAGATTTACATTTGTTAGTTTCATAGAAA 182
Db	1179 CCTTCAATAAAAAATTACAATAACCAACTCCATATAAAAAATTTTCTATTTACGATAAAAAAT 112
QY	183 TCAATTTCTAGAGTCATAAATAATGCCTAAATTTCTATTTCTATTTTCCCTTAAGAA 242
Db	1119 CCAATAAATACACGAATTTTTTCTCTTATTTTCTTTATTCACCTCTTTTCAATTCAA 106
QY	243 AAAAAACAGCCCATTA 259
Db	1059 ATATTTTCAACGCTCTTAA 1043


```

FT intron /number= 6
FT 4386..4694
FT /*tag= 1
FT /number= 6
FT 4695..4909
FT /*tag= m
FT /number= 7
FT 4910..5014
FT /*tag= n
FT /number= 7
FT 5015..5198
FT /*tag= o
FT /number= 8
FT 5199..5300
FT /*tag= p
FT /number= 8
FT 5301..5610
FT /*tag= q
FT /number= 9
XX WO9907223-Al.
XX 18-FEB-1999.
XX 05-AUG-1998; 98WO-US016261.
XX 06-AUG-1997; 97US-00907226.
XX (REGC ) UNIV CALIFORNIA.
XX Etzler ME, Murphy JB;
XX WPI; 1999-167136/14.
XX P-PSDB; AAW85684.
XX New polynucleotides encoding Nod factor binding lectins - useful for
XX production of transgenic plants which are able to fix nitrogen.
XX Disclosure; Page 44-47; 57pp; English.
XX The NBP46 root lectin is instrumental in recognising and binding to
XX nitrogen fixing rhizobial bacteria via a lectin-carbohydrate interaction.
XX The production of transgenic plants comprising an expression cassette
XX expressing the NBP46 root lectin is advantageous since it would mean that
XX non-leguminous plants could fix nitrogen from the atmosphere, lessening
XX the need for the addition of nitrogen containing fertilizer to soil. This
XX would lead to higher crop yields where soil has been overplanted and
XX replenishment of the depleted soil with usable nitrogen. Alternatively,
XX expression of NBP46 can be used to modulate oligosaccharide signalling in
XX the plant. The nucleic acid sequences can be used to inhibit expression
XX of an endogenous gene and also to suppress endogenous NBP46 gene
XX expression
XX
XX Query Match 3.1%; Score 51.4; DB 2; Length 6265;
XX Best Local Similarity 49.4%; Pred. No. 0.037;
XX Matches 133; Conservative 0; Mismatches 136; Indels 0; Gaps 0;
XX
XX 24 TCATCTACATGATGATAAATTTTATAGTAGATTACTTCTGTAAATCAATTCAGAGCATGA 83
XX |||||
XX 3725 TCCGCTTAAAGGTAAATTTTCAGAAAAAATACATATGTAATTAATATATATAT 3666
XX |||||
XX 84 AAAAAAATCGCAAGAACAAATATTTTAAATGAATGATGCAATATACAAATTTAATA 143
XX |||||
XX 3665 AATATAACATAAAATTAATATATTTTATTTATTTATATATATATTAATAAATAAATA 3606
XX |||||
XX 144 CACAAATTATGTAAGATTACATTTTGTAGTTTCATAGAATCAATTTCTAGAGTCATAATA 203
XX |||||
XX 3605 AAATAAATATTAATTAATATTTTATATTTTATTAATTAATAAATAAATAAATAATA 3546
XX |||||
XX 204 ATGCCATAAATTAATTTCTATTTCTTAAAGAAAAAAGCCATTAAGGGA 263
XX |||||

```

```

Db 3545 AAATATAAATTTAAAAATAAATAATTTTATTTCATAAATAAATAAATAAATAATTT 3486
QY 264 CCATTAGATGCGCGTCTCCATTTTT 292
Db 3485 AAAATATAAATAAATAATTTATTCACTTTTT 3457

RESULT 7
ABL32303/c
ID ABL32303 standard; DNA; 12763 BP.
XX
AC ABL32303;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 276.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cyostatic; neutropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007537.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
PS Claim 1; SEQ ID NO 276; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SQ Sequence 12763 BP; 3481 A; 219 C; 3026 G; 6037 T; 0 U; 0 Other;
Query Match 3.1%; Score 50.8; DB 6; Length 12763;
Best Local Similarity 56.6%; Pred. No. 0.072;
Matches 94; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
XX
QY 83 AAAAAAATCGCAAGAACAAATATTTTAAATGAATGATGCAATATACAAATTTAATT 142
XX |||||
XX 2935 AAAAAAATTTAAATTAACACCAATTAATAATAATAATAAATAAATTTTATT 2876
XX |||||
QY 143 ACACAATTATCAGATTACATTTGTTAGTTTCTAGAAATCAATTTCTAGAGTCATAAT 202
XX |||||
Db 2875 TATTATTATATTTTAATAAATACCTTTTAAACAAATTTAAAAAATTTCTTCTATTCC 2816
XX |||||
QY 203 AATGCCTAAATAAATTTCTATTCTATTTTCTTCTTAAAGAAAAA 248

```

Qy

205 TGCCATAAATAAAAATTCATTCTATTTTCTTAAGAAAAAA 246

Dβ

11282 TACAATAATAATATAAAAATTTCTTTTAAATCAAAATTCAAAA 11241

RESULT 9
AAS46772/C
ID AAS46772 standard; DNA; 17280 BP.

XX AAS46772;

XX
DT 18-DEC-2001 (first entry)

XX Tumour suppressor gene derived chemically modified sequence #496.

Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;
tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
cytosine methylation; ds.

PD 20-SEP-2001.

PF 15-MAR-2001; 2001WO-EP002955.

PR 15-MAR-2000; 2000DE-01013847.

PR 07-APR-2000; 2000DE-01019173.

PR 01-SEP-2000; 2000DE-01043826.

PA (EPIG-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

XX

genes and oncogenes, useful in designing primers and probes for analyzing diseases associated with cytosine methylation state e.g. cancer.

2000

bases of a segment of chemically pretreated DNA (CP DNA) e.g. with bisulphite, of genes associated with tumour suppression and oncogenes having a sequence taken from 536 (actually 533 since numbers 408, 458 and 500 are missing from the sequence listing) sequences (Ss) and sequences complementary to (Ss). The nucleic acid may be a peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may form part of a set of probes for detecting the cytosine methylation state and/or single nucleotide polymorphisms and also to be used in an array for analysing diseases associated with CpG dinucleotides e.g. cancers and tumours. The probes can also be used in a method for ascertaining genot and/or

Sequence 113515 BP: 31803 A: 1174 C: 24020 G: 56518 T: 0 U: 0 Other:

Query Match	3.1%	Score 50.8;	DB 6;	Length 113515;
Best Local Similarity	51.8%	Pred. No. 0.18;		
Matches 115:	Conservative	0;	Mismatches 107;	Indels 0;
	Conservative			Gaps 0;

25 CATCTACATGTGATAAAATTTTATAGTTAGATTACTTCTTGTAAATCATTTTCAGAGGATGAA 84

Db 11462 CAAATATATATACAAATATATATACAAATATATATTTAATATACATATA 11403

85 AAAAAAAAAATCGCAAGAAAGCAAAATATTTTAAATGAATGATGCAATATACAAATTTAATTAC 144

db 11402 AAAATAAACAAA TTAATAAATATATTTCTATTATTTTAAATATTCTATTACAATCAC 11343

145 ACAATTATGTAAGATTACATTGTTTAGTTTCATAGAAATCAATTTCTAGAGTCATAATAA 204

11342 AATAAAATAAAA TTTTAAATCATATTC AAAAAAAAAATACCATATTAAAAACACAAAAAAT 11283

Query Match

3.1%; Score 50.6; DB 4; Length 17280;

3.1%; Score 50.6; DB 4; Length 17280;

	Best Local Similarity	52.1%;	Pred. No. 0.091;			
	Matches 113;	Conservative	0;	Mismatches 104;	Indels	Gaps 0;
Qy	38	TAAATTTTATAGTAGAATTACTCTTGTAATCATTTTCAGAGGATCGAAAAATAATTCGAA	97			
Db	5424	TAAATAACAAAATAAATTTTCATCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA				
Qy	98	GAAGAAGCAAATATTTTAAATGAATGATGCATATACAATTTAAATTACACATTTATGTGAAG	157			
Db	5364	AATCTCTATNTAAATAATCTCT	5305			
Qy	158	ATTACATGTGTTAGTTTTCATAGAAATCAATTTCTAGAGTCATAATAATGCCATAATTTAA	217			
Db	5304	AAAAACATAAATACTCTTCACCTTAATAATTTTTACCATTTCTTAATAATCCTTATCTAAAT	5245			
Qy	218	ATTCATNTCTATTTTTCCTTAAGAAAAAAAACAGCCC	254			
Db	5244	CAACTAATACACTAAAAAATTAAAAAAAATAAATAATACC	5208			

RESULT 10
ABT07549
ID ABT07549 standard; cDNA; 1134 BP.
XX
AC ABT07549;
XX
DT 14-NOV-2002 (first entry)
XX
DE Human breast cancer associated coding sequence SEQ ID NO: 28.
XX
KW Human; breast specific gene; breast specific protein; breast cancer;
KW gene therapy; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200264611-A1.
XX
PD 22-AUG-2002.
XX
PF 12-FEB-2002; 2002WO-US004197.
XX
PR 13-FEB-2001; 2001US-0268292P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferkey R;
PI Sun Y, Liu C;
XX
DR WPI; 2002-657582/70.
XX
PT New breast specific nucleic acids and proteins, useful for identifying,
PT diagnosing, monitoring, staging, imaging, and treating breast cancer and
PT non-cancerous disease states in breast tissue, and in gene therapy.
XX
PS Claim 1; Page 171; 367pp; English.
XX
CC The present invention provides human breast specific coding sequences and
CC proteins. These can be used in the diagnosis and treatment of breast
CC cancer and non-cancerous diseases of the breast. The present sequence is
CC a coding sequence of the invention
SQ Sequence 1134 BP; 418 A; 141 C; 170 G; 405 T; 0 U; 0 Other.

	Query Match	3.0%;	Score 50.4;	DB 6;	Length 1134;
	Best Local Similarity	51.3%;	Pred. No. 0.032;		
	Matches 117;	Conservative 0;	Mismatches 111;	Indels 0;	Gaps 0
QY	11	CTCAAAATACCA	GTCATCTACATG	TGTAATAATTTT	TAGTTAGTACTCTCTTGTAATCA 70
Db	29	CTTGAGTTCTAG	TCAATTTAATTTG	TGTCCTCAATAAGTTT	TCTGCAATCTCCTTTTTTAAGT 88
QY	71	TTTCAGAGGATG	AAAAAATAATTCG	CAAGAAGCAATAATTTT	TAATGAATGATGCAATAT 130

Db	89	TCCAGAAATTCCTATTATAAATAAGTGTACTTTTACCAATTCCTATGTATAGCAAAACAG	148
Qy	131	ACAAATTTAAATTACACAATTATGTAAGATTACATGTTTTAGTTTCATGAATCAATCAATTC	190
Db	149	ACACCTTTTAGAAAGGATAAGTATCATCAATTTGTTTTTAAAAAAAACAAATTC	208
Qy	191	TAGAGTCATAATAATGCGCTAAATTAATAATCTATTCTATTTTTCCTTA	238
Db	209	CAGACTACTAAATTTGGCATAGAAATAATCTTTTAAAAATGCAACATA	256

RESULT 11
ABL33206
ID ABL33206 standard; DNA; 15767 BP.

XX ABL33206:

XX
DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 1179.

Human; immune system disease; cytosine methylation; antiasthmatic;
 antiarteriosclerotic; anianaemic; cystostatic; nootropic;
 neuroprotective; anti-HIV; anticonvulsant; ophthalmologic;
 antirheumatic; antithyroid; antidiabetic; antipsoriatic;
 antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 ds.

XX Homo sapiens.
OS

XX
PN WO200200928-A2XX
PD 03-JAN-2003

2007-11-20 11:20:00

PF 02-JUL-2001; 2001WO-EP007537.
XX

PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.

XX
PA (EPIC-) EPICENOMICS NC

XX (EFFIG-7) EFFIGENOMICS AG

Pl Oleg A, Piepenbrock C, Berlin K;
XX

DR WPI; 2002-130909/17.
XX

PT Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.

XX
PS
Claim 1; SEQ ID NO 1179; 32pp + Sequence Listing; German

xx The present invention provides a number of human immune system associated
cc genes which are modified by the methylation of cytosines. The sequences
cc can be used in the diagnosis and treatment of immune system disorders,
cc including eye diseases such as retinopathy, neovascular glaucoma and
cc macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
cc leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
cc rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
cc diseases. The present sequence is a gene of the invention

SQ Sequence 15767 BP; 4558 A; 123 C; 3384 G; 7702 T; 0 U; 0 Other;

Query Match 3.0%; Score 50.4; DB 6; Length 15767;
Best Local Similarity 49.3%; Pred. No. 0.099;
Matches 122; Consensus 100%;

132; conservative 0; mismatches 136; indels 0; gaps 0

33 TGTGATAAAATTTTATAGTTAGATTACTTCTTGTAATCATTTTCAGAGGATGAAAAAAT 92

D6 14988 TATTTGAGATGTGAAGTTAAATTTAAAAATTTTGATGAGTTTAAAGTTTAAATTTATTTTTT 15047

Qy 93 CGCAAGAAAGCAAATATTTTAAATGAATGATGCAATATACAAATTTAAATTACACAATTAT 152

CC disorders allowing for improved and informed treatment of patients
 XX Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;
 SQ Query Match 3.0%; Score 50.2; DB 7; Length 8056;
 Best Local Similarity 53.3%; Pred. No. 0.083;
 Matches 106; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
 QY 35 TGAATAATTTTACGTAGTACTTCTGTAATCATTTTCAGAGGATGAAAAAATCG 94
 Db 1527 TAAAAAATTTTAAATAAATTAATTTTAAAAAATTAATTTTAAATTTTAAATTTTAA 1586
 QY 95 CAAGAAGCAAAATTTTAAATGAATGATGCAATATACAAATTTAATTTACACAAATATGT 154
 Db 1587 ATTATATTTTAAATTTTAAATTTTAAAAATTTGAATTAATGAATTTGTAATTTAAAA 1646
 QY 155 AGATTACATTTGTTAGTTTCATAGAAATCAATTTCTAGAGTCATATTAATGCTAAAT 214
 Db 1647 AAAATTTTAAATTAATATGATAAAATTTTATTTTAAATTTTAAATTTTAAATTTTAA 1706
 QY 215 AAAATTTCTATTTCTATTTT 233
 Db 1707 AATATAATTAATATATTT 1725
 RESULT 14
 ID AAS45498/c
 XX AAS45498 standard; DNA; 17142 BP.
 AC AAS45498;
 DT 18-DEC-2001 (first entry)
 DE Chemically pretreated genomic DNA associated with cell cycle #102.
 KW Cell cycle; human; CpG dinucleotide; cytosine methylation; HIV; aging;
 KW human immunodeficiency virus; neurodegenerative disorder; solid tumour;
 KW graft-versus-host disease; glomerular disease; Lewy body disease; cancer;
 KW arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;
 KW immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;
 KW PCR primer.
 OS Homo sapiens.
 PN WO200168911-A2.
 XX 20-SEP-2001.
 XX 15-MAR-2001; 2001WO-EP002945.
 XX 15-MAR-2000; 2000DE-01013847.
 PR 06-APR-2000; 2000DE-01019058.
 PR 07-APR-2000; 2000DE-01019173.
 PR 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX (EPIG-) EPIGENOMICS AG.
 PI Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2001-602751/68.
 DR Designing primers and probes for analyzing diseases associated with
 PT cytosine methylation state e.g. arthritis, cancer, aging,
 PT arteriosclerosis comprising fragments of chemically modified genes
 PT associated with cell cycle.
 XX Claim 1; SEQ ID NO 203; 28pp; English.
 CC Sequences AAS45296-AAS45520 represent chemically pretreated genomic DNA
 CC molecules associated with the cell cycle and specific PCR primers of the
 CC invention. The sequences are useful for detecting the methylation state
 CC of all CpG dinucleotides in a sequence and therefore for analysing

CC associated diseases. By analysing cytosine methylations in the pretreated
 CC DNA, genetic and/or epigenetic parameters for the diagnosis and therapy
 CC of existing diseases or the predisposition to specific diseases can be
 CC ascertained. The parameters may be compared to another set of genetic
 CC and/or epigenetic parameters, the differences serving as basis for
 CC diagnosis and/or prognosis events which are disadvantageous to patients.
 CC The sequences of the invention are useful for the diagnosis and therapy
 CC of HIV infection, neurodegenerative disorders, graft-versus-host disease,
 CC aging, glomerular disease, Lewy body disease, arthritis,
 CC arteriosclerosis, solid tumours and cancers
 XX Sequence 17142 BP; 5054 A; 299 C; 3741 G; 8048 T; 0 U; 0 Other;
 SQ Query Match 3.0%; Score 50.2; DB 4; Length 17142;
 Best Local Similarity 50.6%; Pred. No. 0.11;
 Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
 QY 17 ATACCACTCATCTACATGTGATAAATTTTATAGTTAGTACTTCTGTAATCATTTTCAG 76
 Db 9510 ATATCACAATATCTATAAATTAACCTTTTAAAAACCAAACTTTTATTCAAAAATCTCCA 9451
 QY 77 AGGATCAAAAAAATTCGCAAGAACCAATATTTTAAATGAATGATGCAATATACAAAT 136
 Db 9450 CCTATAATTTAAAACTTATAAATACTAACTAACTAACTAACTAACTAACTAACTAA 9391
 QY 137 TTAATTACACAATTAATGTAAGATTACATTTGTTAGTTTCATAGAAATCAATTTCTAGAGT 196
 Db 9390 AAAAACTTTAAAAATTTAAATATCTACCACTAACTAACTAACTAACTAACTAACTAA 9331
 QY 197 CATATAATGCTTAATTAATAATTTCTATTTCTATTTTCTTAAAGAAAAAAGAGCCCA 255
 Db 9330 CTTAATTTATATTAATAAATAATTTTCCAAAAATAATTAAGTAAAAAATAATATCA 9272
 RESULT 15
 ID ABL34107/c
 XX ABL34107 standard; DNA; 17142 BP.
 AC ABL34107;
 DT 26-MAR-2002 (first entry)
 DE Human immune system associated gene SEQ ID NO: 2080.
 KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 ds.
 OS Homo sapiens.
 PN WO200200928-A2.
 XX 03-JAN-2002.
 XX 02-JUL-2001; 2001WO-EP007537.
 PR 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX (EPIG-) EPIGENOMICS AG.
 PI Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2002-130909/17.
 DR Nucleic acid comprising fragment of chemically modified gene, useful for
 PT diagnosis and treatment of diseases associated with abnormal cytosine
 PT methylation.

us-10-732-721-1.rng

XX	Claim 1; SEQ ID NO 2080; 32pp + Sequence Listing; German.
PS	
XX	
CC	The present invention provides a number of human immune system associated
CC	genes which are modified by the methylation of cytosines. The sequences
CC	can be used in the diagnosis and treatment of immune system disorders,
CC	including eye diseases such as retinopathy, neovascular glaucoma and
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC	leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC	diseases. The present sequence is a gene of the invention
XX	
SQ	Sequence 17142 BP; 5054 A; 299 C; 3741 G; 8048 T; 0 U; 0 Other;
	Query Match 3.0%; Score 50.2; DB 6; Length 17142;
	Best Local Similarity 50.6%; Pred. No. 0.11; Indels 0; Gaps 0;
	Matches 121; Conservative 0; Mismatches 118;
Qy	17 ATACCAGCTCATCTACATGTGATAAATTTTATAGTTAGATTACTTCTTGTTAAATCAATTCAG 76
Db	9510 ATATCAAAATATCTATAAAAATACTTTAAAAACAACAACTTTTATTCCAAATATCTCCA 9451
Qy	77 AGGATGAIAAAAAAAAAATCGCAGAAGCAATATTTTAAATCAATGCATATACAAT 136
Db	9450 CCTATAATTAATAAACCTTATAAAAAAATCTAAACCTTAAATCTAACCTTCATTCACAAA 9391
Qy	137 TTAATTTACACAATATGTAAGATTACATTTGTTTAGTTTCAPAGAAATCAATTTCTAGAGT 196
Db	9390 AAAAAGCTAAAAATTTAAATATCTACCAATACTAATTTATTAATACTTTTTTAAATTT 9331
Qy	197 CATATAATAGCCTAAATTTAAATTTCTATCTATTTTCCITTAAGAAAAAACAAGCCCA 255
Db	9330 CTTAATTTATATATTAATAAAAAATTTTCCAAAAATTAATTCGPTAAAAAATAAAAAATATCA 9272

Search completed: September 2, 2004, 07:52:48
Job time : 643 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 07:31:08 ; Search time 144 Seconds
(without alignments)
6389.643 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattcagggtcacataac.....ctcgttagttgggagcgcg 1658

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A-COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B-COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A-COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B-COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS-COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	53.6	3.2	2750	3	US-08-617-860B-33
C 2	51.4	3.1	6265	4	US-09-129-112-3
C 3	48.8	2.9	8961	4	US-10-204-708-80
C 4	46.4	2.8	2448	4	US-09-134-001C-777
C 5	46.4	2.8	9347	4	US-10-204-708-35
C 6	46	2.8	20674	4	US-09-641-638-651
C 7	44.8	2.7	5852	1	US-07-867-106-2
C 8	44	2.7	6656	4	US-10-204-708-75
C 9	44	2.7	168575	4	US-09-426-230-1
C 10	43.8	2.6	5152	4	US-10-204-708-74
C 11	43.8	2.6	1664976	4	US-08-916-421B-1
C 12	43.4	2.6	6182	4	US-10-204-708-87
C 13	43	2.6	1296	4	US-09-107-532A-3402
C 14	43	2.6	1296	4	US-09-134-000C-3111
C 15	43	2.6	5852	1	US-07-867-106-2
C 16	43	2.6	10640	4	US-09-417-485D-5
C 17	42.6	2.6	1850	3	US-08-617-860B-32
C 18	42.6	2.6	4098	2	US-08-605-106-4
C 19	42.6	2.6	9636	1	US-08-323-170B-1
C 20	42.6	2.6	9636	4	US-08-954-441-1
C 21	42.6	2.6	51952	3	US-08-947-823-1
C 22	42.4	2.6	11049	4	US-10-204-708-22
C 23	42.2	2.5	19025	4	US-09-849-334-3
C 24	42.2	2.5	19025	4	US-10-274-878-3
C 25	42	2.5	565	4	US-08-956-171E-1007
C 26	42	2.5	6669	4	US-10-204-708-6
C 27	41.8	2.5	11015	4	US-10-204-708-56

C 28	41.8	2.5	53332	4	US-09-801-861-3	Sequence 3, Appli
C 29	41.6	2.5	1422	1	US-08-319-704-5	Sequence 5, Appli
C 30	41.6	2.5	5562	4	US-10-204-708-63	Sequence 63, Appli
C 31	41.4	2.5	6152	3	US-08-973-462-1	Sequence 1, Appli
C 32	41.4	2.5	640681	4	US-09-790-988-1	Sequence 1, Appli
C 33	41.4	2.5	1664976	4	US-08-916-421B-1	Sequence 1, Appli
C 34	41	2.5	19124	2	US-08-487-826B-13	Sequence 13, Appli
C 35	40.8	2.5	672	4	US-09-134-001C-1992	Sequence 1992, Ap
C 36	40.8	2.5	6070	4	US-10-204-708-10	Sequence 10, Appli
C 37	40.8	2.5	580073	4	US-08-545-528D-1	Sequence 1, Appli
C 38	40.6	2.4	1784	4	US-09-601-198-21	Sequence 21, Appli
C 39	40.4	2.4	357	4	US-09-328-352-538	Sequence 538, App
C 40	40.4	2.4	1450	3	US-08-617-860B-22	Sequence 22, Appli
C 41	40.4	2.4	6768	1	US-08-107-755A-1	Sequence 1, Appli
C 42	40.4	2.4	8093	4	US-10-204-708-31	Sequence 31, Appli
C 43	40.4	2.4	8457	1	US-07-991-867B-1	Sequence 1, Appli
C 44	40.4	2.4	8457	2	US-08-544-332-1	Sequence 1, Appli
C 45	40.4	2.4	8457	4	US-09-370-861A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-617-860B-33/C
; Sequence 33, Application US/08617860B
; Patent No. 6133506
; GENERAL INFORMATION:
; APPLICANT: Typfer, R., Bautor, J., Bothmann, H., Filsak, E.,
; APPLICANT: Hvricke-Grandpierre, C., Klein, B., Martini, N.,
; APPLICANT: Mller, A., Schulte, W., Voetz, M., Walek, J.,
; APPLICANT: Schell, J.
; TITLE OF INVENTION: Promoters
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steinberg, Raskin & Davidson, P.C.
; STREET: 1140 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617,860B
; FILING DATE: 01-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/02950
; FILING DATE: 05-SEP-1994
; APPLICATION NUMBER: DE P4329951.2
; FILING DATE: 04-SEP-1993
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2750 Base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: Double stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Cuphea lanceolata
; IMMEDIATE SOURCE:
; LIBRARY: genomic Lambda FIX II
; CLONE: CITE94
; FEATURE:
; NAME/KEY: Startcodon
; LOCATION: 2637..2639
; FEATURE:
; NAME/KEY: CDS

QY	144	CACAATTATGTAAGATTACATTGTTTGTAGTTTTCATAGAAATCAATTTCTCAGGTCAATAA	203
Db	3605	AAATATATATTTTAAATAATTTATATTTTATATAATAATAAATAAATAAATAAATAA	3546
QY	204	ATGCTAAATTTAAATTTCTATTTCTTTTCCCTTTAGAAAAAACAAGCCCATTAAGGGA	263
Db	3545	AAATATAAATTTAAAAATAAATAATTTTATTTTTCATATAATAATAATAATAATAAT	3486
QY	264	CAATTAGATGGCGCTCTCCATTTT	292
Db	3485	AAATATAAATAATAATTAATCACTTTT	3457
RESULT 3			
US-10-204-708-80/c			
; Sequence 80, Application US/10204708			
; Patent No. 6677731			
; GENERAL INFORMATION:			
; APPLICANT: OLEK, Alexander			
; APPLICANT: PIEPENBROCK, Christian			
; APPLICANT: BERLIN, Kurt			
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication			
; TITLE OF INVENTION: By Assessing DNA Methylation			
; FILE REFERENCE: 5013.1012			
; CURRENT APPLICATION NUMBER: US/10/204,708			
; CURRENT FILING DATE: 2003-05-06			
; PRIOR APPLICATION NUMBER: PCT/EP01/03971			
; PRIOR FILING DATE: 2001-04-06			
; PRIOR APPLICATION NUMBER: DE 10019058.8			
; PRIOR FILING DATE: 2000-04-06			
; PRIOR APPLICATION NUMBER: DE 10019173.8			
; PRIOR FILING DATE: 2000-04-07			
; PRIOR APPLICATION NUMBER: DE 10032529.7			
; PRIOR FILING DATE: 2000-06-30			
; PRIOR APPLICATION NUMBER: DE 10043826.1			
; PRIOR FILING DATE: 2000-09-01			
; NUMBER OF SEQ ID NOS: 98			
; SEQ ID NO 80			
; LENGTH: 8961			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)			
; FEATURE:			
; NAME/KEY: unsure			
; LOCATION: (5096)			
; OTHER INFORMATION: n is a or g c o r t			
US-10-204-708-80			
Query Match 2.9%; Score 48.8; DB 4; Length 8961;			
Best Local Similarity 51.4%; Pred. No. 0.0076;			
Matches 113; Conservative 0; Mismatches 107; Indels 0; Gaps 0			
QY	33	TGTGATAAATTTTATAGTTAGATTACTTCTTGTATCAATTTCCAGAGGATGAAAAAAT	92
Db	5927	TATAATATATTTAAATTAACACGAACCTCTTCAATTTATATAATATCAAAATATAA	5866
QY	93	CGCAAGAAAGCAAAATATTTTAAATGAATGATCAATATACAAATTTTAAATTTACAAATTTAT	152
Db	5867	TAAAAAATAAATTTATATAAATAAATTAACCTTTTCAATTTTAAAAAATAAATAAAT	5808
QY	153	GTAAGATTACATTTGTTTATAGTTTTCATAGAAATCAATTTCTAGAGTCATATAATGCTAAA	212
Db	5807	CCATCTATTAACACGAATAAATACTATTATTAATCAATTTATTAATTTCTTTTAAATAA	5748
QY	213	TTAAAAATTTCTATTCTATTTTTCTTAAAGAAAAAACAAC	252
Db	5747	AAATACTTACTATAAATCTATAAATAAATAAATAAATAA	5708
RESULT 4			
US-09-134-001C-777			

; Sequence 777, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 777
; LENGTH: 2448
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-777

Query Match 2.8%; Score 46.4; DB 4; Length 2448;
Best Local Similarity 52.0%; Pred. No. 0.017;
Matches 104; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 26 ATCTACATGTCGATAAATTTTATAGTTAGATTACTTCTGTAATCATTTTCAGAGGATGAAA 85
Db 743 ATCTACCTCGTACATTTGTTATGAAACACACAAAATTTCTCAATGAGATATTGAGA 802
QY 86 AAAAAATCGAAGAAGCAAAATATTTTAAATGAATGATGCAATATACAAATTTAATTACA 145
Db 803 AACCAATTGAAACAGAGCAATTTACTAGAAATCATTTGACGTTTTATCAGAAGAGAAAAT 862
QY 146 CAATATGATGATACATTTGTTTGTAGTTCTATAGAAATCAATTTCTAGAGTCATTAAT 205
Db 863 TAATATCTGAAGCTGATCAGGTATCAATTTCCAAAGTTTACTATTCTAGAAATGAAAAGTG 922
QY 206 GCCTAAATTTAAATTTCTATT 225
Db 923 TGCAAACTTATACCGAATT 942

RESULT 5
US-10-204-708-35/c
; Sequence 35, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 35
; LENGTH: 9347
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-35

Query Match 2.8%; Score 46.4; DB 4; Length 9347;
Best Local Similarity 54.8%; Pred. No. 0.035;
Matches 92; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 83 AAAAAAATCGAAGAAGCAAAATATTTTAAATGAATGATGCAATATACAAATTTAATT 142
Db 7686 AAAAAAATACTAAATAAATAAATAAATAAATCAACAAATATTTTAAAAACAAATTA 7627
QY 143 ACACAATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 202
Db 7626 ACACATTTCTA 7567
QY 203 AATGCTTAAATTAATTTCTATTTCTATTTCTTCTTCTTCTTCTTCTTCTTCTTCT 250
Db 7566 AAAAAAATACTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7519

RESULT 6
US-09-641-638-651/c
; Sequence 651, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 651
; LENGTH: 20674
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: 1123..3123
; OTHER INFORMATION: 5' regulatory region
; NAME/KEY: exon
; LOCATION: 3124..3297
; OTHER INFORMATION: exon 1
; NAME/KEY: exon
; LOCATION: 3871..4072
; OTHER INFORMATION: exon 2
; NAME/KEY: exon
; LOCATION: 5552..5633
; OTHER INFORMATION: exon 3
; NAME/KEY: exon
; LOCATION: 5758..5880
; OTHER INFORMATION: exon 4
; NAME/KEY: exon
; LOCATION: 5996..6099
; OTHER INFORMATION: exon 5
; NAME/KEY: exon
; LOCATION: 6349..6509
; OTHER INFORMATION: exon 6
; NAME/KEY: exon
; LOCATION: 7379..7522
; OTHER INFORMATION: exon 7
; NAME/KEY: exon
; LOCATION: 8645..8854
; OTHER INFORMATION: exon 8

```

; NAME/KEY: exon
; LOCATION: 12254..12340
; OTHER INFORMATION: exon 9
; NAME/KEY: exon
; LOCATION: 12854..13023
; OTHER INFORMATION: exon 10
; NAME/KEY: exon
; LOCATION: 13308..13429
; OTHER INFORMATION: exon 11
; NAME/KEY: exon
; LOCATION: 16567..16667
; OTHER INFORMATION: exon 12
; NAME/KEY: exon
; LOCATION: 16775..16945
; OTHER INFORMATION: exon 13
; NAME/KEY: exon
; LOCATION: 17063..17554
; OTHER INFORMATION: exon 14
; NAME/KEY: misc feature
; LOCATION: 17555..20674
; OTHER INFORMATION: 3'regulatory region
; NAME/KEY: allele
; LOCATION: 1128
; OTHER INFORMATION: 10-508-191 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 1182
; OTHER INFORMATION: 10-508-245 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 1559
; OTHER INFORMATION: 10-509-284 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 1570
; OTHER INFORMATION: 10-509-295 : deletion of C
; NAME/KEY: allele
; LOCATION: 1827
; OTHER INFORMATION: 10-510-173 : variable motif ATTTA or TTTTTT
; NAME/KEY: allele
; LOCATION: 2048
; OTHER INFORMATION: 10-511-62 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 2323
; OTHER INFORMATION: 10-511-337 : insertion of T
; NAME/KEY: allele
; LOCATION: 2341
; OTHER INFORMATION: 10-512-36 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 2623
; OTHER INFORMATION: 10-512-318 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 2832
; OTHER INFORMATION: 10-513-250 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 2844
; OTHER INFORMATION: 10-513-262 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 2934
; OTHER INFORMATION: 10-513-352 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 2947
; OTHER INFORMATION: 10-513-365 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 3802
; OTHER INFORMATION: 12-206-81 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 4062
; OTHER INFORMATION: 10-343-231 : deletion of C
; NAME/KEY: allele
; LOCATION: 4088
; OTHER INFORMATION: 12-206-366 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 4109
; OTHER INFORMATION: 10-343-278 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 4170
; OTHER INFORMATION: 10-343-339 : polymorphic base G or T
; NAME/KEY: allele
; LOCATION: 5903
; OTHER INFORMATION: 10-346-23 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 6019
; OTHER INFORMATION: 10-346-141 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 6141
; OTHER INFORMATION: 10-346-263 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 6183
; OTHER INFORMATION: 10-346-305 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 6338
; OTHER INFORMATION: 10-347-74 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 6375
; OTHER INFORMATION: 10-347-111 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 6429
; OTHER INFORMATION: 10-347-165 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 6467
; OTHER INFORMATION: 10-347-203 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 6484
; OTHER INFORMATION: 10-347-220 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 6534
; OTHER INFORMATION: 10-347-271 : polymorphic base A or T
; NAME/KEY: allele
; LOCATION: 6611
; OTHER INFORMATION: 10-347-348 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 7668
; OTHER INFORMATION: 10-348-391 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 8608
; OTHER INFORMATION: 10-349-47 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 8658
; OTHER INFORMATION: 10-349-97 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 8703
; OTHER INFORMATION: 10-349-142 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 8777
; OTHER INFORMATION: 10-349-216 : deletion of CTG
; NAME/KEY: allele
; LOCATION: 8785
; OTHER INFORMATION: 10-349-224 : polymorphic base G or T
; NAME/KEY: allele
; LOCATION: 8926
; OTHER INFORMATION: 10-349-368 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 12171
; OTHER INFORMATION: 10-350-72 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 12429
; OTHER INFORMATION: 10-350-332 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 13341
; OTHER INFORMATION: 10-507-170 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 13492
; OTHER INFORMATION: 10-507-321 : polymorphic base A or C
; NAME/KEY: allele
; LOCATION: 13524
; OTHER INFORMATION: 10-507-353 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 13535

```


Db 1138 AATATAATTAAATCTATTACTCACCAATCTTTCAATATCTAAAT 1091

RESULT 9

US-09-426-290-1/c
; Sequence 1, Application US/09426290
; Patent No. 6410712
; GENERAL INFORMATION:
; APPLICANT: Berglind Ran Olafsdottir
; APPLICANT: Jeffrey Gulcher
; TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
; FILE REFERENCE: 2345,2001-000
; CURRENT APPLICATION NUMBER: US/09/426,290
; CURRENT FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1

LENGTH: 168575

TYPE: DNA

ORGANISM: Homo Sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (21181)...(21403)

NAME/KEY: CDS

LOCATION: (95252)...(95430)

NAME/KEY: CDS

LOCATION: (101753)...(101996)

NAME/KEY: CDS

LOCATION: (110324)...(110439)

NAME/KEY: CDS

LOCATION: (124058)...(124278)

NAME/KEY: CDS

LOCATION: (127009)...(127130)

NAME/KEY: CDS

LOCATION: (128910)...(129139)

US-09-426-290-1

Query Match 2.7%; Score 44; DB 4; Length 168575;

Best Local Similarity 47.8%; Pred. No. 0.74; Indels 0; Gaps 0;
Matches 128; Conservative 0; Mismatches 140;

QY 29 TACATGTGATAAATTTATAGTTAGATTACTCTTGTAAATCAATTCAGAGGATGAAAAA 88
Db 109154 TATATATATATATATATATATATAGTTATATATATATATATATATATATATAT 109095
QY 89 AATCCCAAGAAAGCAATATTTTAAATGAATGCAATATACAAATTTAATTACAA 148
Db 109094 ATATATTTAGATATAATAAATAATATATATATATATATATATATATATATAT 109035
QY 149 TTATGTAAGATTACATTTGTTTTCATAGAAATCAATTTCTAGAGTCATAATAATGCC 208
Db 109034 ATAAAT 108975
QY 209 TAAATTAATCTTATCTATTTTCTTAAAGAAAAAAGCCCATTAAGGACCAT 268
Db 108974 TATTTAT 108915
QY 269 AGAATGGCGCTCTCCATTTTATATA 296
Db 108914 TGAATTTTGTATGTATATTTTATTA 108887

RESULT 10

US-10-204-708-74/c
; Sequence 74, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012

; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 74
; LENGTH: 5152
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-74

Query Match 2.6%; Score 43.8; DB 4; Length 5152;
Best Local Similarity 51.8%; Pred. No. 0.13; Indels 0; Gaps 0;
Matches 99; Conservative 0; Mismatches 92;

QY 83 AAAAAAATCGCAAGAAAGCAATATTTTAAATGAATGCAATATACAAATTTAAAT 142
Db 4914 AAATTAATTTCTTTTACAAAAATATATATATATATATATATATATATATAT 4855
QY 143 ACACAATTAATGAAGATTACATTTTGTAGTTTCATAGAAATCAATTTCTAGAGTCATAT 202
Db 4854 AAATTAACATATATTTTAAACAATAAATACCTTATTTAAACTTATTTACTCATTTATTC 4795
QY 203 AATGCTTAATTAATTTCTATCTATTTTCTTAAAGAAAAAAGCCCATTAAGGG 262
Db 4794 TATACTTAACATAACCTCATTTATAATATATACCTTAAAAAATTAATATCCATAAAC 4735
QY 263 ACCATTAGAAT 273
Db 4734 AAATTAATAT 4724

RESULT 11

US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature

```
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g

LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
```

US-08-916-421B-1

Query Match 2.6%; Score 43.8; DB 4; Length 1664976;
 Best Local Similarity 47.0%; Pred. No. 2.9;
 Matches 135; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

QY 35 TGATAAATTTTAGTTAGTACTTCTTGTAATCATTTTCAGAGATGAAAAAATCG 94
 DB 159284 TAAACATACAAATAAACAACAACTCATAAATCTCTTAAATAAATAAATG 159343

QY 95 CAAGAAAGCAATATTTTAAATGAATGCAATATACAAATTTTAAATACACAATATGT 154
 DB 159344 AAAAATAGTAATCTTTTATTATTTTCCAAATACCAAAATCAAAACCTTACTTATAA 159403

QY 155 AAGATTACATTTTGTAGTTTCAGAAATCAATTTTCAGAGTCAATTAATGCGCTAAAT 214
 DB 159404 TCTTAAAAATCGAAAGATTTCTTAAACCTGTTGCTATGCTCAACAAGAAGCAAAAT 159463

QY 215 AAAATTTCTATTCTTTTCTTAAAGAAAAAAGAGCCCATTAAGGGACCAATTAGAAATG 274
 DB 159464 AAAATTAATAAATCTATTATGCAATTAATAATTTCTCAATAAGCAATAATCTATTATTT 159523

QY 275 CCGCTGCTCCATTTTTTATATATAGATATGAGTTGTTGCTCAC 321
 DB 159524 TATACATCACTATTGTCATTAATGATAATGATAAATTACTGCTGAC 159570

RESULT 12
 US-10-204-708-87/c
 ; Sequence 87, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; TITLE OF INVENTION: by Assessing DNA Methylation
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; CURRENT FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/039971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 87
 ; LENGTH: 6182
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-204-708-87

Query Match 2.6%; Score 43.4; DB 4; Length 6182;
 Best Local Similarity 48.6%; Pred. No. 0.18;
 Matches 119; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 2 AATTCAGGCTCACATACACGTCATCTACATGATAAATTTTATAGTTAGTACTTC 61
 DB 3691 AAAACACAACTTTAAATTTAAAAAACCTTAAAAAATAAATCAAAATTTTCTTTA 3632

QY 62 TTGTAATCATTTTCAGAGGATGAAAAAATCGCAAGCAAGCAATATTTTAAATGAATG 121
 DB 3631 CAAAAAATTTTCGGATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3572

QY 122 ATGCAATATACAAATTTAATACAAATTTATGTAAGATACATTTGTTTAGTTTCATAGAA 181

Db 3571 ACCTAAAAAACAACAAATACGTTTTTACATTTTAAAAATTTACTATTATATAA 3512

QY 182 ATCAATTTCTAGAGTCATAAATGCGCTAAATTAATAATCTATTTCTATTTTCCCTAAGA 241
 DB 3511 ATACATTTCTTAATAAATAAATAAATCAATTTTAAATACACATTTCTCACTTTTACT 3452

QY 242 AAAAA 246
 DB 3451 AATAA 3447

RESULT 13
 US-09-107-532A-3402
 ; Sequence 3402, Application US/09107532A
 ; Patent No. 6593275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02154
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD-ROM ISO9660
 ; COMPUTER: PC
 ; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,532A
 ; FILING DATE: 30-Jun-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/085,598
 ; FILING DATE: 14 May 1998
 ; APPLICATION NUMBER: 60/051571
 ; FILING DATE: July 2, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ariniello, Pamela Deneke
 ; REGISTRATION NUMBER: 40,489
 ; REFERENCE/DOCKET NUMBER: GTC-012
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (781)893-5007
 ; TELEFAX: (781)893-8277
 ; INFORMATION FOR SEQ ID NO: 3402:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1296 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: circular
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Enterococcus faecium
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (B) LOCATION 1...1296
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 3402:
 US-09-107-532A-3402

Query Match 2.6%; Score 43; DB 4; Length 1296;
 Best Local Similarity 49.3%; Pred. No. 0.099;
 Matches 112; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 41 ATTATATAGTTAGTACTTCTTGTATCATTTTCAGAGGATGAAAAAATCGCAAGAA 100
 DB 408 ATTAGTAGTGGAGTATTTACAAATTAAGAAGTAAAAATATAGAAGAAATGATAAGAA 467

QY 101 AGCAAAATTTTAAATGAATGCAATATACAAATTTTAAATACAAATTTATGTAGATT 160

Db 468 TTCAATTATTTGTCATGAAGATTGCTTAACAAACAACTCTAAATTAGGTGATGAAGT 527
QY 161 ACAITGTTAGTTTCATAGAAATCAATTTCTAGAGTCATATAATGCGCTAAATTTAAAT 220
Db 528 TGATCTTGAATTTACTAGATATTTGAAAGAAAGTGGAAGAAATATAAGTCATAAATTTAAAT 587
QY 221 CTATTCATTTTTCCTTAAGAAACAAACAGCCCATTAAGGACCAT 267
Db 588 TATAGGAATCTTTCTGTAAAAACAGAAACATATACAGGATTAT 634

RESULT 14

US-09-134-000C-3111
; Sequence 3111, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3111
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3111

Query Match 2.6%; Score 43; DB 4; Length 1296;
Best Local Similarity 49.3%; Pred. No. 0.099;
Matches 112; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 41 ATTTTATAGTTAGATTACTCTTGTATCATTTTCAGAGGATGAAAAAATTCGAGAA 100
Db 408 ATTATAGTGGAGATTTCATTAATTAAGAAGGTAAATATAGAGAAATGATAGAA 467
QY 101 AGCAATATTTTAAATGAATGATGCAATATACAAATTTTAAATACACAAATTTGTAAGATT 160
Db 468 TTCAATTTATTTCTCATGAAGATTGCTTAACAAACAACTCTAAATTTAGGTGATGAAGT 527
QY 161 ACAITGTTAGTTTCATAGAAATCAATTTCTAGAGTCATATAATGCGCTAAATTTAAAT 220
Db 528 TGATCTTGAATTTACTAGATATTTGAAAGAAAGTGGAAGAAATATAAGTCATAAATTTAAAT 587
QY 221 CTATTCATTTTTCCTTAAGAAACAAACAGCCCATTAAGGACCAT 267
Db 588 TATAGGAATCTTTCTGTAAAAACAGAAACATATACAGGATTAT 634

RESULT 15

US-07-867-106-2/c
; Sequence 2, Application US/07867106
; Patent No. 5389526
; GENERAL INFORMATION:
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/867,106
FILING DATE: 19920625
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7187
APPLICATION NUMBER: PCT/AU90/00530
FILING DATE: 02-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Feeney, Joanne Longo
REGISTRATION NUMBER: 35,134
REFERENCE/DOCKET NUMBER: RICE-0002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5852 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
US-07-867-106-2

Query Match 2.6%; Score 43; DB 1; Length 5852;
Best Local Similarity 54.8%; Pred. No. 0.22;
Matches 85; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 83 AAAAAAATCGCAAGAAAGCAATATTTTAAATGAATGATGCAATATACAAATTTAAT 142
Db 1850 AAGAAAAAATAAAAAAAGTAGAATTTATTAAAAATTTAAAAATTTTCAATCTTAATA 1791
QY 143 ACACAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 202
Db 1790 AATTAAGTATATATCGATAGCAATTTATTTTTATATCTATCTAAAAAATACTAGGAA 1731
QY 203 AATGCCATAATTTAAATTTCTATTCTTTTCTCTT 237
Db 1730 AATGAATGTCATCAATAGTATTTTAAATTTT 1696

Search completed: September 2, 2004, 10:46:19
Job time : 150 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 09:38:09 ; Search time 785 Seconds
(without alignments)
10498.497 Million cell updates/sec

Title: US-10-732-721-1

Perfect score: 1658

Sequence: 1 gaattcaggctcaataac.....ctcgttagttggagcgcg 1658

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3267054 seqs, 2485319735 residues

Total number of hits satisfying chosen parameters: 6534108

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1658	100.0	1658	17	US-10-732-721-1
2	98.4	5.9	836	13	Sequence 1, Appli
3	97.4	5.9	918	13	Sequence 7252, Ap
4	71.4	4.3	894	13	Sequence 13779, A
5	53.2	3.2	7498	15	Sequence 3295, Ap
6	52.2	3.1	6012	13	Sequence 230, App
7	51.4	3.1	6265	9	Sequence 236, App
8	50.8	3.1	12763	15	Sequence 3, Appli
9	50.8	3.1	113515	15	Sequence 276, App
10	50.6	3.1	17280	13	Sequence 2147, Ap
11	50.4	3.0	1134	15	Sequence 498, App
12	50.4	3.0	15767	15	Sequence 28, Appl
13	50.4	3.0	15767	15	Sequence 1179, Ap
14	50.2	3.0	17142	15	Sequence 105, App
					Sequence 206, App

C 15	50.2	3.0	17142	15	US-10-311-455-2080	Sequence 2080, Ap
C 16	50.2	3.0	17142	15	US-10-240-453-304	Sequence 304, App
C 17	49.6	3.0	653	13	US-10-027-632-104828	Sequence 104828,
C 18	49.6	3.0	653	16	US-10-027-632-104828	Sequence 104828,
C 19	49.6	3.0	778	13	US-10-027-632-142028	Sequence 142028,
C 20	49.6	3.0	778	16	US-10-027-632-142028	Sequence 142028,
C 21	49.6	3.0	3673778	15	US-10-312-841-1	Sequence 1, Appli
C 22	49.4	3.0	591	17	US-10-437-963-21828	Sequence 21828, A
C 23	49.4	3.0	14551	15	US-10-240-485-137	Sequence 137, App
C 24	48.8	2.9	8961	15	US-10-204-708-80	Sequence 80, Appl
C 25	48.8	2.9	8961	15	US-10-240-453-302	Sequence 302, App
C 26	48.6	2.9	5891	13	US-10-221-613-43	Sequence 43, Appl
C 27	48.4	2.9	632	9	US-09-772-134B-57	Sequence 57, Appl
C 28	48.4	2.9	736	9	US-09-772-134B-53	Sequence 53, Appl
C 29	48.2	2.9	5313	15	US-10-311-455-736	Sequence 736, App
C 30	48.2	2.9	5768	15	US-10-311-455-2136	Sequence 2136, Ap
C 31	48.2	2.9	7276	15	US-10-311-455-875	Sequence 875, App
C 32	48.2	2.9	13038	15	US-10-311-455-1248	Sequence 1248, Ap
C 33	48.2	2.9	3673778	15	US-10-312-841-2	Sequence 2, Appli
C 34	48	2.9	6053	15	US-10-239-676-76	Sequence 76, Appl
C 35	48	2.9	6053	15	US-10-240-453-82	Sequence 82, Appl
C 36	48	2.9	7657	15	US-10-239-676-185	Sequence 185, App
C 37	48	2.9	7657	15	US-10-311-455-1995	Sequence 1995, Ap
C 38	48	2.9	33053	17	US-10-433-793-36	Sequence 36, Appl
C 39	47.8	2.9	3673778	15	US-10-312-841-1	Sequence 1, Appli
C 40	47.6	2.9	5499	17	US-10-433-793-2	Sequence 2, Appli
C 41	47.6	2.9	9963	15	US-10-311-455-667	Sequence 667, App
C 42	47.6	2.9	18997	15	US-10-172-086-17	Sequence 17, Appl
C 43	47.6	2.9	18997	15	US-10-311-455-543	Sequence 543, App
C 44	47.6	2.9	18997	17	US-10-311-507-31	Sequence 31, Appl
C 45	47.6	2.9	113515	15	US-10-311-455-2148	Sequence 2148, Ap

ALIGNMENTS

RESULT 1

US-10-732-721-1
; Sequence 1, Application US/10732721
; Publication No. US20040163144A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 38-15(52826)A
; CURRENT APPLICATION NUMBER: US/10/732,721
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/60/434,242
; PRIOR FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1658
; TYPE: DNA
; ORGANISM: Zea mays
US-10-732-721-1

Query Match 100.0%; Score 1658; DB 17; Length 1658;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1658; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GAATTCAGGCTCAATACAGTCATCATGTGATAAATTTTAGTTAGTACTT	60
Db	1	GAATTCAGGCTCAATACAGTCATCATGTGATAAATTTTAGTTAGTACTT	60
Qy	61	CTTGTAAATCATTTTCAGAGGATGAAAAAATCCCAAGAAAGCAATATTTTAATGAAT	120
Db	61	CTTGTAAATCATTTTCAGAGGATGAAAAAATCCCAAGAAAGCAATATTTTAATGAAT	120
Qy	121	GATCCAATATACAAATTTAATTAATGACAAATTTATGTAAGATTACATTTTTCATAGA	180
Db	121	GATCCAATATACAAATTTAATTAATGACAAATTTATGTAAGATTACATTTTTCATAGA	180


```

; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 13779
; LENGTH: 918
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB143-042-D8_FLI
US-10-425-114-13779

Query Match      5.9%; Score 97.4; DB 13; Length 918;
Best Local Similarity 99.0%; Pred. No. 1.4e-13;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1554 CAGCAACAGCCCAACACCTAAACTAAAGAACACAGTAGTACCTCCCTGTGTACGTCTAGCAACG 1613
Db      1 CAGCAACAGCCCAACACCTAAACTAAAGAACACAGTAGTACCTCCCTGTGTACGTCTAGCAACG 60

QY 1614 ATCCGCCGCCAGCTCGCCAGTTTTGCTCGTTAGTTTGGG 1652
Db      61 ATCCGCCGCCAGCTCGCCAGTTTTGCTCGTTAGTTTGGG 99

RESULT 4
US-10-425-114-3295
; Sequence 3295, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3295
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700258323_FLI
US-10-425-114-3295

Query Match      4.3%; Score 71.4; DB 13; Length 894;
Best Local Similarity 98.6%; Pred. No. 4.4e-07;
Matches 72; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1580 GAACAGTAGTAGTCCCTGTGTACGCTAGCAACGATCCGCCGCCAGCTCGCCAGTTTGC 1639
Db      1 GAACAGTAGTAGTCCCTGTGTACGCTAGCAACGATCCGCCGCCAGCTCGCCAGTTTGC 60

QY 1640 TCGTTAGTTTGGG 1652
Db      61 TCGTTAGTTTGGG 73

RESULT 5
US-10-311-455-230/c
; Sequence 230, Application US/10311455
; Publication No. US20030143606A1

```

	Best Local Similarity	49.4%;	Pred. No. 0.16;	Mismatches 136;	Indels 0;	Gaps 0;
	Matches 133;	Conservative	0;			
QY	24	TCATCTACATGTCATATAAATTTATAGTTAGATTACCTTCCTGTAATCATTTTCAGAGATGA	83			
Db	3725	TCGCGTTAAAGGTAAAAATTTTCAGAAAAAANAATACATATGTTATTAATTAATTTATAT	3666			
QY	84	AAAAAATAATCGCAAGAAAGCAATAATTTTTAAATGAATGATGCAATATACAAATTTTAATTA	143			
Db	3665	AATATAACATAAATAATTAATTTTTATTATTATTATATTATTAATTAATAATAAATAATTA	3606			
QY	144	CACAANTATGTAAGATTACATCTGTTTAGTTTCATAGAAATCAATTTCTAGAGTCATAATA	203			
Db	3605	AATAATTAATTATTAATAATTTATATTTTATATAATAAATAAATAAATAATAATA	3546			
QY	204	ATGCCATAATTAATAATTCATATCTATTTTTTCCTTAAGAAAAAAGCCCATTAAGGGA	263			
Db	3545	AAATATAAATTTAAAAATAAATAATTTATTTTTCATATAATAAATAAATAAATAATTT	3486			
QY	264	CCATTAGAATGCGCGTGTGCCATTTTTT	292			
Db	3485	AAATATAAATAAATAATTAATTCACACTTTTT	3457			

```

RESULT E
US-10-31
; Sequen
; Public
; GENERA
; APPLJ
; APPLJ
; APPLJ
; TITLE
; FILE
; CURR
; PRIOR
; PRIOR
; PRIOR
; PRIOR
; PRIOR
; NUMB
; SEQ
; LEN
; TYPE
; ORG
; FEAT
; OTH
US-10-31

```

Query Match	3.1%	Score 50.8;	DB 15;	Length 12763;
Best Local Similarity	56.6%;	Pred. No. 0.38;		
Matches 94;	Conservative 0;	Mismatches 72;	Indels 0;	Gaps 0
Qy	83	AAAAAAAAATCGCAAGAAGCAAAATATTTTAAATGAATGATGCAATATACAAAATTTAAT	142	
Db	2935	AAAAAAAAATTTAAATTTAAATTTAAACCACTTAAATATAATATTAATAAAAAATTTTATT	2876	
Qy	143	ACACAATTATGTAAGATTACATTTCTTTAGTTTCATAGAAATCAATTTCTAGAGTCATAAT	202	
Db	2875	TATTTATTATATTTTAAATAAATACCTTTTAAACAAATTTAAAAAATTTCTTTCTATTCTT	2816	
Qy	203	AATGCCTAAATTTAAAAATCTATTCTATTCTTCTTAAAGAAAAAAA	248	
Db	2815	AATTTATAAATAAAAAATAATTT	2770	

RESULT 9

US-10-311-455-2147/c

; Sequence 2147, Application US/10311455

; Publication No. US20030143606A1

```
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2003-12-16
; PRIOR FILING DATE: 2001-07-02
; PRIOR FILING DATE: 2001-07-02
; PRIOR FILING DATE: 2000-06-30
; PRIOR FILING DATE: 2000-09-01
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2147
; LENGTH: 113515
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2147

Query Match          3.1%; Score 50.8; DB 15; Length 113515;
Best Local Similarity 51.8%; Pred. No. 1.7; Indels 0; Gaps 0;
Matches 115; Conservative 0; Mismatches 107;

QY 25 CATCTACATGTCATAAATTTTATAGTTAGTACTCTTCTGTAATCATTTTCAGAGGATCAA 84
DB 11462 CAAATATATATACAAATATATATATACAAATATATATATTTAATATCTACATATA 11403

QY 85 AAAAAAATCGAAGAACCAATATTTTAAATGAATGCAATATACAAATTTTAATTAC 144
DB 11402 AAAATAACCAATTAATAATATATATTTCTATTATTTTAAATATCTTATTAACAATCAC 11343

QY 145 ACAATTATGTAAGATTACATTTGTTAGTTTTCATAGAAATCAATTTCTAGAGTCATAATA 204
DB 11342 AATAAATAAATTTTAATCATATTCAAAAAATAATACCATATTTAAACACAAAAAAT 11283

QY 205 TGCCTAAATTAATTCATTCTATTCTTTTCCTTAAGAAAAAA 246
DB 11282 TACAATAAATAATAAATTTCTTTTAAATCAAAATTCAAA 11241

RESULT 10
US-10-221-714A-498/c
; Sequence 498, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; FILE OF INVENTION: tumor suppressor genes and oncogenes
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221,714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR FILING DATE: 2001-03-15
; PRIOR FILING DATE: 2001-03-15
; PRIOR FILING DATE: 2000-03-15
; PRIOR FILING DATE: 2000-03-15
; PRIOR FILING DATE: 2000-04-06
; PRIOR FILING DATE: 2000-04-06
; PRIOR FILING DATE: 2000-04-07
; PRIOR FILING DATE: 2000-04-07
; PRIOR FILING DATE: 2000-05-29.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 498
; LENGTH: 17280
```

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-498

Query Match          3.1%; Score 50.6; DB 13; Length 17280;
Best Local Similarity 52.1%; Pred. No. 0.52;
Matches 113; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 38 TAAATTTTATAGTTAGTACTCTTCTGTAATCATTTTCAGAGGATGAAAAAATCGCAA 97
DB 5424 TAAATAACAAATATAATTTTCATCTCAAAAAAATAAATAATCTATATATAAATAACTCT 5365

QY 98 GAAAGCAAAATATTTTAAATGAATGATGCAATATACAAATTTTAAATTTACAAATTTATGTAA 157
DB 5364 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5305

QY 158 ATTACATTTGTTAGTTTCATAGAATCAATTTCTAGAGTCATAATAATGCCTAAATTTAA 217
DB 5304 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5245

QY 218 ATTCTATTCTATTTTCTCTTAAGAAAAAACAAGCC 254
DB 5244 CACTAATACACTAAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5208

RESULT 11
US-10-074-475-28
; Sequence 28, Application US/10074475
; Publication No. US20030092898A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Karra, Kalpana
; APPLICANT: Caferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; FILE OF INVENTION: Genes and Proteins
; FILE REFERENCE: DEX-0313
; CURRENT APPLICATION NUMBER: US/10/074,475
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: 60/268,292
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 295
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-074-475-28

Query Match          3.0%; Score 50.4; DB 15; Length 1134;
Best Local Similarity 51.3%; Pred. No. 0.091;
Matches 117; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 11 CTCACAATACCAGTCATCTACATGATGATAAATTTTATAGTTAGATTACTCTCTGTAATCA 70
DB 29 CTGGAAGTCTAGTCAATTTTAATTTGATCCAAATGTTTCTGAAATCTCTCTTTTAA 88

QY 71 TTTCCAGAGGATGAAAAAATAATCGAAGAACCAAAATATTTTAAATGAATGATCAATAT 130
DB 89 TCCAAGAAATTTCTATTATAAATAAGTGACTTTTACCAATTTCCATTTGATATAAGCAACAG 148

QY 131 ACAATTTAATTAACAAATTAATGTAAGTACATTTGTTAGTTTCATAGAAATCAATTTTC 190
DB 149 ACACCTTTTAGAAAAGGATAAGTAATCATCAATTTGTTTGTATTTTAAAAAATAAATAATTC 208

QY 191 TAGAGTCATAATAATGCCTAAATTAATAATCTATTCTATTCTTTCTTA 238
```

Db 209 CAGACTACTAAATTTGGCATAAGAAATAATCTTTTAAATGCAACATA 256

RESULT 12

US-10-311-455-1179
; Sequence 1179, Application US/10311455
; Publication No. US20030143606A1

GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424

SEQ ID NO 1179

LENGTH: 15767

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-1179

Query Match 3.0%; Score 50.4; DB 15; Length 15767;

Best Local Similarity 49.3%; Pred. No. 0.55;

Matches 132; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 33 TGTGATAAAATTTTATAGTTAGTACTTCTTGTAAATCAATTCAGAGGATGAAAAAAAT 92

Db 14988 TATTTTGAGATGGAAGTTAAATTTTAAATTTTGAAGTTTAAAGTTTAAATTTATTTT 15047

QY 93 CGCAAGAAACAAATATTTTAAATGAATGATGCAATATACAAATTTAATACAAATAT 152

Db 15048 TGTAAAGAAATATTTTATAGATAATGTTGTATATTTTATATTTTATATACGAAAGTAT 15107

QY 153 GTAAGATTACATTTGTTAGTTTTCATAGAAATCAATTTCTAGAGTCAATAATGCGCTAAA 212

Db 15108 GTAAATGTTGATTTTATTTTGGTGAATTTTAAAGTCAATTTGTTGTTTAAA 15167

QY 213 TTAATAATCTATCTATTTTCCCTTAAGAAAAAAGAGCCCATTAAGGACCATTAGAA 272

Db 15168 TTGAATTTAAATATTTTATTTTAAATTTAAATTTGTAATTTGTAATTTAAGTGTATTAGAG 15227

QY 273 TGGCGCTGCTCCATTTTATATATA 300

Db 15228 AAGTTTATATTTATTTTGTGTTTTTTA 15255

RESULT 13

US-10-240-485-105
; Sequence 105, Application US/10240485
; Publication No. US20030148327A1

GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240,485
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06

Query Match 3.0%; Score 50.2; DB 15; Length 17142;

Best Local Similarity 50.6%; Pred. No. 0.65;

; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202

SEQ ID NO 105

LENGTH: 15767

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-240-485-105

Query Match 3.0%; Score 50.4; DB 15; Length 15767;

Best Local Similarity 49.3%; Pred. No. 0.55;

Matches 132; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 33 TGTGATAAAATTTTATAGTTAGTACTTCTTGTAAATCAATTCAGAGGATGAAAAAAAT 92

Db 14988 TATTTTGAGATGGAAGTTAAATTTTAAATTTTGAAGTTTAAAGTTTAAATTTATTTT 15047

QY 93 CGCAAGAAACAAATATTTTAAATGAATGATGCAATATACAAATTTAATACAAATAT 152

Db 15048 TGTAAAGAAATATTTTATAGATAATGTTGTATATTTTATATTTTATATACGAAAGTAT 15107

QY 153 GTAAGATTACATTTGTTAGTTTTCATAGAAATCAATTTCTAGAGTCAATAATGCGCTAAA 212

Db 15108 GTAAATGTTGATTTTATTTTGGTGAATTTTAAAGTCAATTTGTTGTTTAAA 15167

QY 213 TTAATAATCTATCTATTTTCCCTTAAGAAAAAAGAGCCCATTAAGGACCATTAGAA 272

Db 15168 TTGAATTTAAATATTTTATTTTAAATTTAAATTTGTAATTTGTAATTTAAGTGTATTAGAG 15227

QY 273 TGGCGCTGCTCCATTTTATATATA 300

Db 15228 AAGTTTATATTTATTTTGTGTTTTTTA 15255

RESULT 14

US-10-239-676-206/c
; Sequence 206, Application US/10239676
; Publication No. US20030082609A1

GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228

SEQ ID NO 206

LENGTH: 17142

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-239-676-206

Query Match 3.0%; Score 50.2; DB 15; Length 17142;

Best Local Similarity 50.6%; Pred. No. 0.65;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 07:41:59 : Search time 4952 Seconds
(without alignments)
11780.813 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattcaggctcaataac.....ctcgttagtttgggacggcg 1658

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA Main:*

- 1: /cgn2_6/ptodata/2/pna/PCRTUS COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06 COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07 COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US080 COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US081 COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US082 COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US083 COMB.seq.*
- 8: /cgn2_6/ptodata/2/pna/US084 COMB.seq.*
- 9: /cgn2_6/ptodata/2/pna/US085 COMB.seq.*
- 10: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
- 11: /cgn2_6/ptodata/2/pna/US087 COMB.seq.*
- 12: /cgn2_6/ptodata/2/pna/US088 COMB.seq.*
- 13: /cgn2_6/ptodata/2/pna/US089 COMB.seq.*
- 14: /cgn2_6/ptodata/2/pna/US090 COMB.seq.*
- 15: /cgn2_6/ptodata/2/pna/US091 COMB.seq.*
- 16: /cgn2_6/ptodata/2/pna/US092A COMB.seq.*
- 17: /cgn2_6/ptodata/2/pna/US092B COMB.seq.*
- 18: /cgn2_6/ptodata/2/pna/US093A COMB.seq.*
- 19: /cgn2_6/ptodata/2/pna/US093B COMB.seq.*
- 20: /cgn2_6/ptodata/2/pna/US094 COMB.seq.*
- 21: /cgn2_6/ptodata/2/pna/US095A COMB.seq.*
- 22: /cgn2_6/ptodata/2/pna/US095B COMB.seq.*
- 23: /cgn2_6/ptodata/2/pna/US095C COMB.seq.*
- 24: /cgn2_6/ptodata/2/pna/US095D COMB.seq.*
- 25: /cgn2_6/ptodata/2/pna/US096A COMB.seq.*
- 26: /cgn2_6/ptodata/2/pna/US096B COMB.seq.*
- 27: /cgn2_6/ptodata/2/pna/US096C COMB.seq.*
- 28: /cgn2_6/ptodata/2/pna/US096D COMB.seq.*
- 29: /cgn2_6/ptodata/2/pna/US096E COMB.seq.*
- 30: /cgn2_6/ptodata/2/pna/US097A COMB.seq.*
- 31: /cgn2_6/ptodata/2/pna/US097B COMB.seq.*
- 32: /cgn2_6/ptodata/2/pna/US097C COMB.seq.*
- 33: /cgn2_6/ptodata/2/pna/US098A COMB.seq.*
- 34: /cgn2_6/ptodata/2/pna/US098B COMB.seq.*
- 35: /cgn2_6/ptodata/2/pna/US098C COMB.seq.*
- 36: /cgn2_6/ptodata/2/pna/US098D COMB.seq.*
- 37: /cgn2_6/ptodata/2/pna/US099A COMB.seq.*
- 38: /cgn2_6/ptodata/2/pna/US099B COMB.seq.*
- 39: /cgn2_6/ptodata/2/pna/US099C COMB.seq.*
- 40: /cgn2_6/ptodata/2/pna/US099D COMB.seq.*
- 41: /cgn2_6/ptodata/2/pna/US099E COMB.seq.*
- 42: /cgn2_6/ptodata/2/pna/US099F COMB.seq.*
- 43: /cgn2_6/ptodata/2/pna/US100A COMB.seq.*

- 44: /cgn2_6/ptodata/2/pna/US100B COMB.seq.*
- 45: /cgn2_6/ptodata/2/pna/US101A COMB.seq.*
- 46: /cgn2_6/ptodata/2/pna/US101B COMB.seq.*
- 47: /cgn2_6/ptodata/2/pna/US102A COMB.seq.*
- 48: /cgn2_6/ptodata/2/pna/US102B COMB.seq.*
- 49: /cgn2_6/ptodata/2/pna/US103A COMB.seq.*
- 50: /cgn2_6/ptodata/2/pna/US103B COMB.seq.*
- 51: /cgn2_6/ptodata/2/pna/US104A COMB.seq.*
- 52: /cgn2_6/ptodata/2/pna/US104B COMB.seq.*
- 53: /cgn2_6/ptodata/2/pna/US106 COMB.seq.*
- 54: /cgn2_6/ptodata/2/pna/US107A COMB.seq.*
- 55: /cgn2_6/ptodata/2/pna/US107B COMB.seq.*
- 56: /cgn2_6/ptodata/2/pna/US6000 COMB.seq.*
- 57: /cgn2_6/ptodata/2/pna/US6001 COMB.seq.*
- 58: /cgn2_6/ptodata/2/pna/US6002 COMB.seq.*
- 59: /cgn2_6/ptodata/2/pna/US6003 COMB.seq.*
- 60: /cgn2_6/ptodata/2/pna/US6004 COMB.seq.*
- 61: /cgn2_6/ptodata/2/pna/US6005 COMB.seq.*
- 62: /cgn2_6/ptodata/2/pna/US6006 COMB.seq.*
- 63: /cgn2_6/ptodata/2/pna/US6007 COMB.seq.*
- 64: /cgn2_6/ptodata/2/pna/US6008 COMB.seq.*
- 65: /cgn2_6/ptodata/2/pna/US6009 COMB.seq.*
- 66: /cgn2_6/ptodata/2/pna/US6010 COMB.seq.*
- 67: /cgn2_6/ptodata/2/pna/US6011 COMB.seq.*
- 68: /cgn2_6/ptodata/2/pna/US6012 COMB.seq.*
- 69: /cgn2_6/ptodata/2/pna/US6013 COMB.seq.*
- 70: /cgn2_6/ptodata/2/pna/US6014 COMB.seq.*
- 71: /cgn2_6/ptodata/2/pna/US6015 COMB.seq.*
- 72: /cgn2_6/ptodata/2/pna/US6016 COMB.seq.*
- 73: /cgn2_6/ptodata/2/pna/US6017 COMB.seq.*
- 74: /cgn2_6/ptodata/2/pna/US6018 COMB.seq.*
- 75: /cgn2_6/ptodata/2/pna/US6019 COMB.seq.*
- 76: /cgn2_6/ptodata/2/pna/US6020 COMB.seq.*
- 77: /cgn2_6/ptodata/2/pna/US6021 COMB.seq.*
- 78: /cgn2_6/ptodata/2/pna/US6022 COMB.seq.*
- 79: /cgn2_6/ptodata/2/pna/US6023A COMB.seq.*
- 80: /cgn2_6/ptodata/2/pna/US6023B COMB.seq.*
- 81: /cgn2_6/ptodata/2/pna/US6024 COMB.seq.*
- 82: /cgn2_6/ptodata/2/pna/US6025 COMB.seq.*
- 83: /cgn2_6/ptodata/2/pna/US6026 COMB.seq.*
- 84: /cgn2_6/ptodata/2/pna/US6027 COMB.seq.*
- 85: /cgn2_6/ptodata/2/pna/US6028 COMB.seq.*
- 86: /cgn2_6/ptodata/2/pna/US6029 COMB.seq.*
- 87: /cgn2_6/ptodata/2/pna/US6030 COMB.seq.*
- 88: /cgn2_6/ptodata/2/pna/US6031 COMB.seq.*
- 89: /cgn2_6/ptodata/2/pna/US6032 COMB.seq.*
- 90: /cgn2_6/ptodata/2/pna/US6033 COMB.seq.*
- 91: /cgn2_6/ptodata/2/pna/US6034 COMB.seq.*
- 92: /cgn2_6/ptodata/2/pna/US6035 COMB.seq.*
- 93: /cgn2_6/ptodata/2/pna/US6036 COMB.seq.*
- 94: /cgn2_6/ptodata/2/pna/US6037 COMB.seq.*
- 95: /cgn2_6/ptodata/2/pna/US6038 COMB.seq.*
- 96: /cgn2_6/ptodata/2/pna/US6039 COMB.seq.*
- 97: /cgn2_6/ptodata/2/pna/US6040 COMB.seq.*
- 98: /cgn2_6/ptodata/2/pna/US6041 COMB.seq.*
- 99: /cgn2_6/ptodata/2/pna/US6042 COMB.seq.*
- 100: /cgn2_6/ptodata/2/pna/US6043 COMB.seq.*
- 101: /cgn2_6/ptodata/2/pna/US6044 COMB.seq.*
- 102: /cgn2_6/ptodata/2/pna/US6045 COMB.seq.*
- 103: /cgn2_6/ptodata/2/pna/US6046 COMB.seq.*
- 104: /cgn2_6/ptodata/2/pna/US6047 COMB.seq.*
- 105: /cgn2_6/ptodata/2/pna/US6048 COMB.seq.*
- 106: /cgn2_6/ptodata/2/pna/US6049 COMB.seq.*
- 107: /cgn2_6/ptodata/2/pna/US6050 COMB.seq.*
- 108: /cgn2_6/ptodata/2/pna/US6051 COMB.seq.*
- 109: /cgn2_6/ptodata/2/pna/US6052 COMB.seq.*
- 110: /cgn2_6/ptodata/2/pna/US6053 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	Match	Length	DB	ID	Description
1	1658	100.0	100.0	1658	100	US-60-434-242-1	Sequence 1, Appl
2	1658	100.0	100.0	1658	110	US-60-531-039-16	Sequence 16, Appl
3	110.4	6.7	385	26	US-09-620-111B-7324	Sequence 7324, Ap	
4	99.4	6.0	308	18	US-09-304-517A-40406	Sequence 40406, A	
5	99.4	6.0	308	19	US-09-371-146A-40406	Sequence 40406, A	
6	99.4	6.0	308	42	US-09-985-678-40406	Sequence 40406, A	
7	98.4	5.9	396	35	US-09-865-439A-30469	Sequence 30469, A	
8	98.4	5.9	396	76	US-60-207-458-74766	Sequence 74766, A	
9	98.4	5.9	470	35	US-09-865-439A-71134	Sequence 71134, A	
10	98.4	5.9	470	76	US-60-207-458-71431	Sequence 71431, A	
11	98.4	5.9	836	46	US-10-155-881-37309	Sequence 37309, A	
12	98.4	5.9	836	51	US-10-425-114-7252	Sequence 7252, Ap	
13	98.4	5.9	836	51	US-10-425-114A-7252	Sequence 7252, Ap	
14	98.4	5.9	836	88	US-60-312-544-973	Sequence 973, App	
15	97.4	5.9	861	27	US-09-654-617-273716	Sequence 273716, A	
16	97.4	5.9	861	29	US-09-684-016-273716	Sequence 273716, A	
17	97.4	5.9	918	51	US-10-425-114-13779	Sequence 13779, A	
18	97.4	5.9	918	51	US-10-425-114A-13779	Sequence 13779, A	
19	96.8	5.8	315	18	US-09-304-517A-78600	Sequence 78600, A	
20	96.8	5.8	315	19	US-09-371-146A-78600	Sequence 78600, A	
21	96.8	5.8	315	24	US-09-565-306-3420	Sequence 3420, Ap	
22	96.8	5.8	315	42	US-09-985-678-78600	Sequence 78600, A	
23	93.8	5.7	204	18	US-09-304-517A-39577	Sequence 39577, A	
24	93.8	5.7	204	19	US-09-371-146A-39577	Sequence 39577, A	
25	93.8	5.7	204	42	US-09-985-678-39577	Sequence 39577, A	
26	91.4	5.5	527	35	US-09-865-439A-48132	Sequence 48132, A	
27	91.4	5.5	527	76	US-60-207-458-92395	Sequence 92395, A	
28	90.4	5.5	280	18	US-09-303-031A-3994	Sequence 3994, Ap	
29	90.4	5.5	280	18	US-09-304-517A-57027	Sequence 57027, A	
30	90.4	5.5	280	19	US-09-371-146A-57027	Sequence 57027, A	
31	90.4	5.5	280	36	US-09-894-949-3994	Sequence 3994, Ap	
32	90.4	5.5	280	36	US-09-894-949A-3994	Sequence 3994, Ap	
33	90.4	5.5	280	42	US-09-985-678-57027	Sequence 57027, A	
34	83.4	5.0	437	18	US-09-304-517A-79013	Sequence 79013, A	
35	83.4	5.0	437	19	US-09-371-146A-79013	Sequence 79013, A	
36	83.4	5.0	437	24	US-09-565-306-1774	Sequence 1774, Ap	
37	83.4	5.0	437	42	US-09-985-678-79013	Sequence 79013, A	
38	81.4	4.9	460	35	US-09-865-439A-43255	Sequence 43255, A	
39	81.4	4.9	460	76	US-60-207-458-87528	Sequence 87528, A	
40	78.4	4.7	686	35	US-09-865-439A-31938	Sequence 31938, A	
41	78.4	4.7	686	76	US-60-207-458-76235	Sequence 76235, A	
42	73.4	4.4	316	35	US-09-865-439A-44537	Sequence 44537, A	
43	73.4	4.4	316	76	US-60-207-458-88800	Sequence 88800, A	
44	71.4	4.3	291	18	US-09-304-517A-39265	Sequence 39265, A	
45	71.4	4.3	291	19	US-09-371-146A-39265	Sequence 39265, A	

ALIGNMENTS

RESULT 1
 US-60-434-242-1
 ; Sequence 1, Application US/60434242
 ; GENERAL INFORMATION:
 ; APPLICANT: Monsanto Technology, LLC
 ; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
 ; TITLE OF INVENTION: Thereof
 ; FILE REFERENCE: 38-15(52826)A
 ; CURRENT APPLICATION NUMBER: US/60/434,242
 ; CURRENT FILING DATE: 2002-12-18
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 1
 ; LENGTH: 1658
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 US-60-434-242-1
 Query Match 100.0%; Score 1658; DB 100; Length 1658;

Db 1021 TAGAACACAAGAGAGACAAGAGTGTAGAGAGAACTGATCTTTGTACTATATGCTG 1080
QY 1081 CTCTCCAAAGGTTACATGATATGCGGATCTCTCTATTTATAGACAAACTAGGGTTT 1140
Db 1081 CTCTCCAAAGGTTACATGATATGCGGATCTCTCTATTTATAGACAAACTAGGGTTT 1140
QY 1141 CAGGCATATGGGCCACATAGGCTTCTTGCCGCCAAGAAAGGTTCTTAACTACCTACCATCT 1200
Db 1141 CAGGCATATGGGCCACATAGGCTTCTTGCCGCCAAGAAAGGTTCTTAACTACCTACCATCT 1200
QY 1201 TCCAGCGCGGTGCTCGCTGCTTCACTCTCTCCCTCCGCTCCAGAGCGGATAAATACG 1260
Db 1201 TCCAGCGCGGTGCTCGCTGCTTCACTCTCTCCCTCCGCTCCAGAGCGGATAAATACG 1260
QY 1261 GGGTCTACCGTGCACCGCAGCATGTCGCGCGGAGGCGCATCTGCGCGGGGATGCG 1320
Db 1261 GGGTCTACCGTGCACCGCAGCATGTCGCGCGGAGGCGCATCTGCGCGGGGATGCG 1320
QY 1321 CGCAGGCCATCGCTCGACACGACGCTACGTGTGCGCCCGCGAGGCGCTCTCGCACACGC 1380
Db 1321 CGCAGGCCATCGCTCGACACGACGCTACGTGTGCGCCCGCGAGGCGCTCTCGCACACGC 1380
QY 1381 GTCCGCGCCGAGCGCGGAGGACCTTGCAATTCGTCGCGTGGTGCAACGTCCT 1440
Db 1381 GTCCGCGCGCAGCGCGGCGGAGGACCTTGCAATTCGTCGCGTGGTGCAACGTCCT 1440
QY 1441 GGTCCGGCTTGCGGCTTTGCCCTTGCCCTCGCGCGGAGTGTGCGCCCGCGAGGCGCATCTGCGCGGGGATGCG 1500
Db 1441 GGTCCGGCTTGCGGCTTTGCCCTTGCCCTCGCGCGGAGTGTGCGCCCGCGAGGCGCATCTGCGCGGGGATGCG 1500
QY 1501 CGGCTATAAAAGGCGGCGCCACCTGATCTCCATCTCACAAAGCAAGCAGCAGCAAC 1560
Db 1501 CGGCTATAAAAGGCGGCGCCACCTGATCTCCATCTCACAAAGCAAGCAGCAGCAAC 1560
QY 1561 AGCCACACCTTAACCTAAAGACAGTAGTAGTCCCTGTGTAGCTAGCAACGATCCGCG 1620
Db 1561 AGCCACACCTTAACCTAAAGACAGTAGTAGTCCCTGTGTAGCTAGCAACGATCCGCG 1620
QY 1621 CCGAGCTGCCAGTTTGTCTGCTTGTAGTTTGGGACGCGG 1658
Db 1621 CCGAGCTGCCAGTTTGTCTGCTTGTAGTTTGGGACGCGG 1658

RESULT 2
US-60-531-039-16
; Sequence 16, Application US/60531039
; GENERAL INFORMATION:
; APPLICANT: Zheng, Wei
; APPLICANT: Savidge, Beth
; APPLICANT: Savage, Thomas
; APPLICANT: Chaudhuri, Sumita
; TITLE OF INVENTION: Materials and Methods for the Modulation of D-type Cyclins in
; TITLE OF INVENTION: Plants
; FILE REFERENCE: REN-01-083
; CURRENT APPLICATION NUMBER: US/60/531,039
; CURRENT FILING DATE: 2003-12-19
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 1658
; TYPE: DNA
; ORGANISM: Zea mays
US-60-531-039-16

Query Match 100.0%; Score 1658; DB 110; Length 1658;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1658; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAATTCAGGCTCAACATACAGTCACTACATGTGATAAATTTATAGTATTACTT 60
Db 1 GAAATTCAGGCTCAACATACAGTCACTACATGTGATAAATTTATAGTATTACTT 60
QY 61 CTTGTAATCAATTTACAGAGGATGAAAAAATCGCAAGAAAGGTTCTTAAATGAAT 120

Db 61 CTTGTAATCAATTTACAGAGGATGAAAAAATCGCAAGAAAGGTTCTTAAATGAAT 120
QY 121 GATGCAATATACAAATTTAAATTTACAAATTTATGTAAGATTTACATTTGTTAGTTTCATAGA 180
Db 121 GATGCAATATACAAATTTAAATTTACAAATTTATGTAAGATTTACATTTGTTAGTTTCATAGA 180
QY 181 AATCAATTTCTAGAGTCATAAATAGCCCTAAATTTAAATTTCTATTTCTATTTCTCTTAAG 240
Db 181 AATCAATTTCTAGAGTCATAAATAGCCCTAAATTTAAATTTCTATTTCTATTTCTCTTAAG 240
QY 241 AAAAAAAGCAGCCCATTAAGGACCATTAAGATGCGCGCTGCTCCATTTTATATATA 300
Db 241 AAAAAAAGCAGCCCATTAAGGACCATTAAGATGCGCGCTGCTCCATTTTATATATA 300
QY 301 GAGATATAGTTGTTGCTCACTAGGCACTCCACCTTCGCATGCGTCTCTTTATATACATG 360
Db 301 GAGATATAGTTGTTGCTCACTAGGCACTCCACCTTCGCATGCGTCTCTTTATATACATG 360
QY 361 AAAAAATGAGCGCTGCTTTTGCATTAAGGTCCTAGAGATGTACGAGTGCAACGCTTCT 420
Db 361 AAAAAATGAGCGCTGCTTTTGCATTAAGGTCCTAGAGATGTACGAGTGCAACGCTTCT 420
QY 421 ATTCACTGCACGAGTAAATTTTATGCAAGGAGGCGCTTGTATTTCTCAGCAGCT 480
Db 421 ATTCACTGCACGAGTAAATTTTATGCAAGGAGGCGCTTGTATTTCTCAGCAGCT 480
QY 481 CTGCGCGCGCTGCTTCTGTTACGCGAGTGAAGGAGCAGCAGCTCTACGTCGCCACG 540
Db 481 CTGCGCGCGCTGCTTCTGTTACGCGAGTGAAGGAGCAGCAGCTCTACGTCGCCACG 540
QY 541 TGCCTGGCTGCTGCCACCGTGATGGCAAGGCTTGTGCTTGGCTGCTGCTGCTGCTG 600
Db 541 TGCCTGGCTGCTGCCACCGTGATGGCAAGGCTTGTGCTTGGCTGCTGCTGCTGCTG 600
QY 601 AGCGACGAGTTGAAGCTACGCGGCTGCTGCTGCGCGCGCTGCTGCGACGCGAC 660
Db 601 AGCGACGAGTTGAAGCTACGCGGCTGCTGCTGCGCGCGCTGCTGCGACGCGAC 660
QY 661 GACGTCGACGCTGATGCTAGGCTTCTACTTTGTTGTTGTTGTTGTTGTTGTTGTTG 720
Db 661 GACGTCGACGCTGATGCTAGGCTTCTACTTTGTTGTTGTTGTTGTTGTTGTTGTTG 720
QY 721 ATGCAATTAATAACGCTGATCTCATGTGATGATCGACGCTGCTGCGGATTTCAA 780
Db 721 ATGCAATTAATAACGCTGATCTCATGTGATGATCGACGCTGCTGCGGATTTCAA 780
QY 781 ATCAGTAACAATCAATCATGTCATCCCATCGACTTCTGTCGGCATAAATCTGCTAGG 840
Db 781 ATCAGTAACAATCAATCATGTCATCCCATCGACTTCTGTCGGCATAAATCTGCTAGG 840
QY 841 ATTTGTGCTTTTCATGTATGCTTGAATAGAGAGAAAGAAATATATATACATAAAAT 900
Db 841 ATTTGTGCTTTTCATGTATGCTTGAATAGAGAGAAAGAAATATATATACATAAAAT 900
QY 901 GAGTTCAACATCTTTCTTAATACCTTTTGTGTCGATCTCTTTAGTAGAGCCAGTGGTGA 960
Db 901 GAGTTCAACATCTTTCTTAATACCTTTTGTGTCGATCTCTTTAGTAGAGCCAGTGGTGA 960
QY 961 TAACGTTTAAGAAACCCCTTGTACCGAGTGTAGTCCAGGCTGCTTAACCTGAATATAG 1020
Db 961 TAACGTTTAAGAAACCCCTTGTACCGAGTGTAGTCCAGGCTGCTTAACCTGAATATAG 1020
QY 1021 TAGAACACAAGGAGACAGAGTGTAGAGAGAACTGATTTCTTTGTTATATGCTGCTG 1080
Db 1021 TAGAACACAAGGAGACAGAGTGTAGAGAGAACTGATTTCTTTGTTATATGCTGCTG 1080
QY 1081 CTCTCCAAAGGTTACATGATGGGATCTCCTCTTATTTATAGACAAACTAGGGTTT 1140
Db 1081 CTCTCCAAAGGTTACATGATGGGATCTCCTCTTATTTATAGACAAACTAGGGTTT 1140
QY 1141 CAGGCATATGGGCCACATAGGCTTCTTGCCGCCAAGAAAGGTTCTTAACTACCTACCATCT 1200


```
; NUMBER OF SEQ ID NOS: 295529
; SEQ ID NO 40406
; LENGTH: 308
; TYPE: DNA
; ORGANISM: Zea mays
US-09-985-678-40406

Query Match
Best Local Similarity 6.0%; Score 99.4; DB 42; Length 308;
Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1552 AGCAGCAACAGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGTGACGTCTAGCAAA 1611
DB 1 AGCAGCAACAGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGTGACGTCTAGCAAA 60

QY 1612 CGATCCGCGCCGAGCTCGCAGTTTGTCTGTTAGTTGGG 1652
DB 61 CGATCCGCGCCGAGCTCGCAGTTTGTCTGTTAGTTGGG 101

RESULT 7
US-09-865-439A-30469
; Sequence 30469, Application US/09865439A
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51936)B
; CURRENT APPLICATION NUMBER: US/09/865,439A
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,458
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 119126
; SEQ ID NO 30469
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3354-052-Pl-K1-G5
US-09-865-439A-30469

Query Match
Best Local Similarity 5.9%; Score 98.4; DB 35; Length 396;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGTGACGTCTAGCAAC 1612
DB 1 GCAGCAACAGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGTGACGTCTAGCAAC 60

QY 1613 GATCCGCGCCGAGCTCGCAGTTTGTCTGTTAGTTGGG 1652
DB 61 GATCCGCGCCGAGCTCGCAGTTTGTCTGTTAGTTGGG 100

RESULT 8
US-60-207-458-74766
; Sequence 74766, Application US/60207458
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; APPLICANT: Conner, Timothy W.
; APPLICANT: Deikman, Jill
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Laigudi, Raghunath V.
; APPLICANT: Ruan, Yijun G.
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Sammons, R. Douglas
; APPLICANT: Shukla, Hridayabhiranjan
; APPLICANT: Wu, Kunsheng
; APPLICANT: Xu, Nanfei
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
```

```
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 38-21(51936)A
; CURRENT APPLICATION NUMBER: US/60/207,458
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 152403
; SEQ ID NO 74766
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: LIB3354-052-Pl-K1-G5
US-60-207-458-74766

Query Match
Best Local Similarity 5.9%; Score 98.4; DB 76; Length 396;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGTGACGTCTAGCAAC 1612
DB 1 GCAGCAACAGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGTGACGTCTAGCAAC 60

QY 1613 GATCCGCGCCGAGCTCGCAGTTTGTCTGTTAGTTGGG 1652
DB 61 GATCCGCGCCGAGCTCGCAGTTTGTCTGTTAGTTGGG 100

RESULT 9
US-09-865-439A-27134
; Sequence 27134, Application US/09865439A
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51936)B
; CURRENT APPLICATION NUMBER: US/09/865,439A
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,458
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 119126
; SEQ ID NO 27134
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(470)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3354-008-Pl-K1-B11
US-09-865-439A-27134

Query Match
Best Local Similarity 5.9%; Score 98.4; DB 35; Length 470;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGTGACGTCTAGCAAC 1612
DB 1 GCAGCAACAGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGTGACGTCTAGCAAC 60

QY 1613 GATCCGCGCCGAGCTCGCAGTTTGTCTGTTAGTTGGG 1652
DB 61 GATCCGCGCCGAGCTCGCAGTTTGTCTGTTAGTTGGG 100

RESULT 10
US-60-207-458-71431
; Sequence 71431, Application US/60207458
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; APPLICANT: Conner, Timothy W.
; APPLICANT: Deikman, Jill
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
```

; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ruan, Yijun G.
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Sammons, R. Douglas
; APPLICANT: Shukla, Hridayabhiranjan
; APPLICANT: Wu, Kunsheng
; APPLICANT: Xu, Nanfei
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(51936)A
; CURRENT APPLICATION NUMBER: US/60/207,458
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 152403
; SEQ ID NO 71431
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: LIB3354-008-PI-KI-B11
US-60-207-458-71431

Query Match 5.9%; Score 98.4; DB 76; Length 470;
Best Local Similarity 99.0%; Pred. No. 7.8e-12;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1553 GCAGCAACGCGACGCTGCCAGTTTTCGTTAGTTGGG 1652
Db 1 GCAGCAACGCGACGCTGCCAGTTTTCGTTAGTTGGG 100
QY 1613 GATCCGCGCGAGCTGCCAGTTTTCGTTAGTTGGG 1652
Db 61 GATCCGCGCGAGCTGCCAGTTTTCGTTAGTTGGG 100

RESULT 11
US-10-155-881-37309
; Sequence 37309, Application US/10155881
; GENERAL INFORMATION:
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Lutfiyeva, Linda L.
; APPLICANT: McIninch, James
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(15300)J
; CURRENT APPLICATION NUMBER: US/10/155,881
; CURRENT FILING DATE: 2002-05-22
; NUMBER OF SEQ ID NOS: 37595
; SEQ ID NO 37309
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: LIB3354-008-PI-KI-B11
US-10-155-881-37309

Query Match 5.9%; Score 98.4; DB 46; Length 836;
Best Local Similarity 99.0%; Pred. No. 9.8e-12;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1553 GCAGCAACGCGACGCTGCCAGTTTTCGTTAGTTGGG 1652
Db 1 GCAGCAACGCGACGCTGCCAGTTTTCGTTAGTTGGG 100
QY 1613 GATCCGCGCGAGCTGCCAGTTTTCGTTAGTTGGG 1652
Db 61 GATCCGCGCGAGCTGCCAGTTTTCGTTAGTTGGG 100

RESULT 12
US-10-425-114-7252
; Sequence 7252, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7252
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700623908_FLI
US-10-425-114-7252

Query Match 5.9%; Score 98.4; DB 51; Length 836;
Best Local Similarity 99.0%; Pred. No. 9.8e-12;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1553 GCAGCAACGCGACGCTGCCAGTTTTCGTTAGTTGGG 1652
Db 1 GCAGCAACGCGACGCTGCCAGTTTTCGTTAGTTGGG 100
QY 1613 GATCCGCGCGAGCTGCCAGTTTTCGTTAGTTGGG 1652
Db 61 GATCCGCGCGAGCTGCCAGTTTTCGTTAGTTGGG 100

RESULT 13
US-10-425-114A-7252
; Sequence 7252, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7252
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700623908_FLI
US-10-425-114A-7252

Query Match 5.9%; Score 98.4; DB 51; Length 836;
Best Local Similarity 99.0%; Pred. No. 9.8e-12;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1553 GCAGCAACGCGACGCTGCCAGTTTTCGTTAGTTGGG 1652
Db 1 GCAGCAACGCGACGCTGCCAGTTTTCGTTAGTTGGG 100
QY 1613 GATCCGCGCGAGCTGCCAGTTTTCGTTAGTTGGG 1652
Db 61 GATCCGCGCGAGCTGCCAGTTTTCGTTAGTTGGG 100

RESULT 14
US-60-312-544-973
; Sequence 973, Application US/60312544
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei

; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)A
; CURRENT APPLICATION NUMBER: US/60/312,544
; CURRENT FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 10730
; SEQ ID NO 973
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (108)..(449)
; OTHER INFORMATION: Clone ID: 700623908_FLI
US-60-312-544-973

Query Match 5.9%; Score 98.4; DB 88; Length 836;
Best Local Similarity 99.0%; Pred. No. 9.8e-12;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1553 GCAGCAACAGCCAAACACCTAACTAAAGAACAGTAGTCCCTGTGTACGCTAGCAAC 1612
Db 1 GCAGCAACAGCCAAACACCTAACTAAAGAACAGTAGTCCCTGTGTACGCTAGCAAC 60
QY 1613 GATCCGCGCGAGCTCGCCAGTTTGTCTGTTAGTTTGGG 1652
Db 61 GATCCGCGCGAGCTCGCCAGTTTGTCTGTTAGTTTGGG 100

RESULT 15

US-09-654-617-273716
; Sequence 273716, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 273716
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Zea mays
US-09-654-617-273716

Query Match 5.9%; Score 97.4; DB 27; Length 861;
Best Local Similarity 99.0%; Pred. No. 1.7e-11;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1554 CAGCAACAGCCAAACACCTAACTAAAGAACAGTAGTCCCTGTGTACGCTAGCAACG 1613
Db 24 CAGCAACAGCCAAACACCTAACTAAAGAACAGTAGTCCCTGTGTACGCTAGCAACG 83
QY 1614 ATCCGCGCGAGCTCGCCAGTTTGTCTGTTAGTTTGGG 1652
Db 84 ATCCGCGCGAGCTCGCCAGTTTGTCTGTTAGTTTGGG 122

Search completed: September 2, 2004, 12:09:01
Job time : 4955 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 2, 2004, 07:52:53 ; Search time 2370 Seconds
(without alignments)
2423.549 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattcagggtcacatac.....ctcgttagttgggacggcg 1658

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5837357 seqs, 1732150321 residues

Total number of hits satisfying chosen parameters: 11674714

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA New.*
1: /cgn2_6/ptodata/1/pna/PCT NEW COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06 NEW COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07 NEW COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08 NEW COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US09 NEW COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US10 NEW COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US10 NEW COMB.seq.2.*
8: /cgn2_6/ptodata/1/pna/US10 NEW COMB.seq.3.*
9: /cgn2_6/ptodata/1/pna/US60 NEW COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1658	100.0	1658	7	US-10-732-721-1
C 2	109	6.6	868	6	US-10-425-115-68623
C 3	101.6	6.1	528	6	US-10-425-115-47987
C 4	100.4	6.1	975	6	US-10-425-115-181022
C 5	84.8	5.1	2211	6	US-10-425-115-82158
C 6	71.4	4.3	1034	6	US-10-425-115-181033
C 7	55	3.3	533	9	US-60-579-062-32930
C 8	50.6	3.1	17280	6	US-10-221-714B-498
C 9	50.4	3.0	805	9	US-60-579-062-23067
C 10	48.8	2.9	207542	5	US-09-949-003C-4513
C 11	48.8	2.9	207542	6	US-10-893-315-148
C 12	48.8	2.9	207542	6	US-10-902-387-493
C 13	48.8	2.9	207557	5	US-09-949-003C-13956
C 14	48.8	2.9	207557	6	US-10-893-315-134
C 15	48.8	2.9	207557	6	US-10-902-387-465
C 16	48.8	2.9	215927	9	US-60-582-609-19505
C 17	48.8	2.9	228854	9	US-60-548-091-5633
C 18	48.2	2.9	791	5	US-09-404-520B-13329
C 19	48	2.9	753	9	US-60-579-062-40662
C 20	47.2	2.8	1882	1	PCT-US03-41761-40265
C 21	47.2	2.8	1882	1	PCT-US03-41761-40265
C 22	47.2	2.8	1882	1	PCT-US03-41761-40265
C 23	47.2	2.8	6106	6	US-10-221-714B-152
C 24	47	2.8	688	9	US-60-579-062-40503

25	46.4	2.8	2448	6	US-10-902-441-777	Sequence 777, App
26	46.4	2.8	2448	7	US-10-724-972A-360	Sequence 360, App
27	46.2	2.8	916	9	US-60-579-062-24705	Sequence 24705, A
C 28	46.2	2.8	7040	6	US-10-221-714B-162	Sequence 162, App
C 29	46.2	2.8	61020	6	US-10-221-714B-513	Sequence 513, App
C 30	46	2.8	594	9	US-60-579-062-41308	Sequence 41308, A
C 31	45.8	2.8	29993	6	US-10-482-823-3	Sequence 3, Appli
C 32	45.8	2.8	29993	6	US-10-482-823-5	Sequence 5, Appli
C 33	45.8	2.8	38342	6	US-10-221-714B-472	Sequence 472, App
C 34	45.6	2.8	600	1	PCT-US03-41761-20202	Sequence 20202, A
C 35	45.6	2.8	600	1	PCT-US03-41761-20202	Sequence 20202, A
C 36	45.6	2.8	600	1	PCT-US03-41766A-20202	Sequence 20202, A
C 37	45.4	2.7	3952	6	US-10-221-714B-8	Sequence 8, Appli
C 38	45.4	2.7	5371	6	US-10-221-714B-526	Sequence 526, App
C 39	45.4	2.7	7133	6	US-10-221-714B-109	Sequence 109, App
C 40	45.2	2.7	8711	6	US-10-221-714B-423	Sequence 423, App
C 41	45.2	2.7	8845	6	US-10-221-714B-265	Sequence 265, App
C 42	45.2	2.7	8883	6	US-10-221-714B-488	Sequence 488, App
C 43	45.2	2.7	12426	6	US-10-221-714B-203	Sequence 203, App
C 44	45.2	2.7	77504	6	US-10-918-711-3720	Sequence 3720, Ap
C 45	45.2	2.7	77504	6	US-10-918-754-16803	Sequence 16803, A

ALIGNMENTS

RESULT 1
US-10-732-721-1
; Sequence 1, Application US/10732721
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 38-15(52826)A
; CURRENT APPLICATION NUMBER: US/10732,721
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/60/434,242
; PRIOR FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1658
; TYPE: DNA
; ORGANISM: Zea mays
US-10-732-721-1

Query Match		100.0%;	Score 1658;	DB 7;	Length 1658;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1658;		Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1	GAATTCACGGCTCACAATACCAATGTCATCTACATGTGATAAATTTTATAGTTAGATTACTT	60		
Db	1	GAATTCACGGCTCACAATACCAATGTCATCTACATGTGATAAATTTTATAGTTAGATTACTT	60		
Qy	61	CTTGTAATCATTTTCAGAGGTGAAAAAATCGCAAGAACCAATATTTTAAATGAAT	120		
Db	61	CTTGTAATCATTTTCAGAGGTGAAAAAATCGCAAGAACCAATATTTTAAATGAAT	120		
Qy	121	GATGCAATATACAAATTTAATACACAATATATGAAGATTACATTTTATAGTTTCATAGA	180		
Db	121	GATGCAATATACAAATTTAATACACAATATATGAAGATTACATTTTATAGTTTCATAGA	180		
Qy	181	AATCAATTTCTAGAGTCATAATATGCTAAATTTAAATTTCTATTTTCTTTTAAAG	240		
Db	181	AATCAATTTCTAGAGTCATAATATGCTAAATTTAAATTTCTATTTTCTTTTAAAG	240		
Qy	241	AAAAAAGAGCCCATTAAGGACCATTAAGATGCGGCTGCTCCATTTTATATATA	300		
Db	241	AAAAAAGAGCCCATTAAGGACCATTAAGATGCGGCTGCTCCATTTTATATATA	300		
Qy	301	GAGATATGAGTTGTTGCTCACTAGGACCTCCACCTCGCATGGTGTCTTTTATTACATG	360		
Db	301	GAGATATGAGTTGTTGCTCACTAGGACCTCCACCTCGCATGGTGTCTTTTATTACATG	360		

QY 361 AAAAAATGACCGCTGCTTTTGCATTAAGGGTCTAGAGATGACAGTGCACACGCTTCGT 420
 Db |||||
 QY 361 AAAAAATGACCGCTGCTTTTGCATTAAGGGTCTAGAGATGACAGTGCACACGCTTCGT 420
 Db |||||
 QY 421 ATTCACTGACAGTAATTTTTTTTAAAGCAGGAGGCGCTTTGTTTCTCAGCAGCT 480
 Db |||||
 QY 421 ATTCACTGACAGTAATTTTTTTTAAAGCAGGAGGCGCTTTGTTTCTCAGCAGCT 480
 Db |||||
 QY 481 CTGCGCGCCGCTGCTTGTGTAAGGAGTACGAGGAGCAGCAGCTCTACGCTGCCCCAG 540
 Db |||||
 QY 481 CTGCGCGCCGCTGCTTGTGTAAGGAGTACGAGGAGCAGCAGCTCTACGCTGCCCCAG 540
 Db |||||
 QY 541 TGCCTGGCTGCTGCCCCGATGCGCAAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 Db |||||
 QY 601 AGCGACGAGTTGAAGCTACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 Db |||||
 QY 601 AGCGACGAGTTGAAGCTACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 Db |||||
 QY 661 GACGTCGACGCTGATGCTAGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 Db |||||
 QY 721 ATGCACTAAATAAAGTGTGATCTCATGTGATGATGATGATGATGATGATGATGATGATGAT 780
 Db |||||
 QY 721 ATGCACTAAATAAAGTGTGATCTCATGTGATGATGATGATGATGATGATGATGATGATGAT 780
 Db |||||
 QY 781 ATCAGTAACAATCACATGTCATCCCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 Db |||||
 QY 781 ATCAGTAACAATCACATGTCATCCCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 Db |||||
 QY 841 ATTGTGCTTTTCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 Db |||||
 QY 841 ATTGTGCTTTTCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 Db |||||
 QY 901 GAGTCAACATCTTCTTAATACCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 Db |||||
 QY 901 GAGTCAACATCTTCTTAATACCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 Db |||||
 QY 961 TAAGGTTAAGAACCCCTTTTACGAGTGTAGTCAAGGCTGCTTAACTGATGATGATGATGATG 1020
 Db |||||
 QY 961 TAAGGTTAAGAACCCCTTTTACGAGTGTAGTCAAGGCTGCTTAACTGATGATGATGATGATG 1020
 Db |||||
 QY 1021 TAGAACAACAGGAGACAGAGTGTAGAGGAACTGATCTTTTGTACTATATATGCTGCTGCTG 1080
 Db |||||
 QY 1021 TAGAACAACAGGAGACAGAGTGTAGAGGAACTGATCTTTTGTACTATATATGCTGCTGCTG 1080
 Db |||||
 QY 1081 CTCTCCAAAGGTTACATGATATGGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
 Db |||||
 QY 1081 CTCTCCAAAGGTTACATGATATGGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
 Db |||||
 QY 1141 CAGGCATATGCGGCACATAGGCTTCTGCGCCCAAGAAAGGTTCTTAACTACCATCT 1200
 Db |||||
 QY 1141 CAGGCATATGCGGCACATAGGCTTCTGCGCCCAAGAAAGGTTCTTAACTACCATCT 1200
 Db |||||
 QY 1201 TCCAGGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
 Db |||||
 QY 1201 TCCAGGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
 Db |||||
 QY 1261 GGGTCTACGGTGCACCGCGCAGATGCTGCGCGGAGGCTCTCTGCGCGGAGTGGC 1320
 Db |||||
 QY 1261 GGGTCTACGGTGCACCGCGCAGATGCTGCGCGGAGGCTCTCTGCGCGGAGTGGC 1320
 Db |||||
 QY 1321 CGCAGGCGCATCGCTCGACAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
 Db |||||
 QY 1321 CGCAGGCGCATCGCTCGACAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
 Db |||||
 QY 1381 GTCCGCGCGACCGGCGCGAGGACCTTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 Db |||||
 QY 1381 GTCCGCGCGACCGGCGCGAGGACCTTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 Db |||||

QY 1441 GGTCCGGCTTGGCGCTTGGCTGCGCGCGAGCTGCGCGCTCTCCACCGGA 1500
 Db |||||
 QY 1441 GGTCCGGCTTGGCGCTTGGCTGCGCGCGAGCTGCGCGCTCTCCACCGGA 1500
 Db |||||
 QY 1501 CGGCTATATAAAGGCGCGCCACCTGATCTCTCATCTCACAAAGCAAGCAGCAGCAAC 1560
 Db |||||
 QY 1501 CGGCTATATAAAGGCGCGCCACCTGATCTCTCATCTCACAAAGCAAGCAGCAGCAAC 1560
 Db |||||
 QY 1561 AGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
 Db |||||
 QY 1561 AGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
 Db |||||
 QY 1621 CCGAGCTGCCAGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1658
 Db |||||
 QY 1621 CCGAGCTGCCAGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1658
 Db |||||

RESULT 2

US-10-425-115-68623/c
 ; Sequence 68623, Application US/10425115
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 68623
 ; LENGTH: 868
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_162578C.1
 US-10-425-115-68623

Query Match 6.6%; Score 109; DB 6; Length 868;
 Best Local Similarity 74.6%; Pred. No. 2.8e-17;
 Matches 150; Conservative 0; Mismatches 50; Indels 1; Gaps 1;

QY 1 GAATTCACGCTCACAATACCATCATCTACATGATGATGATGATGATGATGATGATGATGATGAT 60
 Db |||||
 QY 556 GAATTCACGCTCACAATACCATCATCTACATGATGATGATGATGATGATGATGATGATGATGAT 497
 Db |||||
 QY 61 CTGTGAATCATTTTACAGGATGAAAAAATCGCAAGAAAGCAAAATATTTTAAATGAAT 120
 Db |||||
 QY 496 ATTGTAATCATTTTACAGGATGAAAAAATCGCAAGAAAGCAAAATATTTTAAATGAAT 438
 Db |||||
 QY 121 GATGCAATATACAAATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
 Db |||||
 QY 437 GGCCCAATATACAAATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 378
 Db |||||
 QY 181 AATCAATTTCTAGAGTCATAA 201
 Db |||||
 QY 377 ATTTCAGAAATAAATTTCTTAA 357
 Db |||||

RESULT 3

US-10-425-115-47987/c
 ; Sequence 47987, Application US/10425115
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28


```
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 47987
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(528)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_14376C.1
US-10-425-115-47987

Query Match          6.1%; Score 101.6; DB 6; Length 528;
Best Local Similarity 91.5%; Pred. No. 1.7e-15;
Matches 118; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 924 CTTTGTCTGATCTCTCTAGTAGACGCGTGTGATAACGTTTAAAGACCCCTTGT 983
Db 129 CCTTGTCTGATCTCTCTAGTAGACGCGTGTGATAACGTTTAAAGACCCCTTGT 70
QY 984 ACCGAGTGTAGTCCAA-GCCTGCTTAACTGAATAATAGTAGAACACAAAGGACAAAGT 1042
Db 69 ACCGAGTGTAGTCCAAAGGGCTCTTANCTGAATAATAGTAGTAATAACAGGACAAAGT 10
QY 1043 GTAGAGAGG 1051
Db 9 GTAGAGAGG 1

RESULT 4
US-10-425-115-181022
; Sequence 181022, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 181022
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_9666C.1
US-10-425-115-181022

Query Match          6.1%; Score 100.4; DB 6; Length 975;
Best Local Similarity 99.0%; Pred. No. 4.5e-15;
Matches 101; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1551 CAGCAGCAACGACCAACCTAACTAAAGACAGTAGTCCCTGTGTAGCTAGCA 1610
Db 21 CAGCAGCAACGACCAACCTAACTAAAGACAGTAGTCCCTGTGTAGCTAGCA 80
QY 1611 ACGATCCGCGCGAGCTCGCCAGTTTTCGTAGTTTGGG 1652
Db 81 ACGATCCGCGCGAGCTCGCCAGTTTTCGTAGTTTGGG 122

RESULT 5
US-10-425-115-82158/c
; Sequence 82158, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 82158
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_174942C.1
US-10-425-115-82158

Query Match          5.1%; Score 84.8; DB 6; Length 2211;
Best Local Similarity 92.6%; Pred. No. 5.9e-11;
Matches 100; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1086 CAAAGGTTACATGATATGGGATCTCTCTCTATTTATAGAC-AAAGTGGTTTCAGG 1144
Db 2211 CCAAGGTTACATGATATGGGATCTCTCTCTATTTATAGAC-AAAGTGGTTTCAGG 2152
QY 1145 CATATGGGCCACATAGGCTTCCTGCGCCCAAGAAAGTTTCTTAACAC 1192
Db 2151 CATATGGGCCACATAGGCTTCCTGCGCCCAAGAAAGTTTCTTAACAC 2104

RESULT 6
US-10-425-115-181033
; Sequence 181033, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 181033
; LENGTH: 1034
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_9667C.1
US-10-425-115-181033

Query Match          4.3%; Score 71.4; DB 6; Length 1034;
Best Local Similarity 98.6%; Pred. No. 1.1e-07;
Matches 72; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1580 GAACAGTAGTAGTCCCTGTGTAGCTAGCAACGATCCGCGCGAGCTCGCCAGTTTTCG 1639
Db 1 GAACAGTAGTAGTCCCTGTGTAGCTAGCAACGATCCGCGCGAGCTCGCCAGTTTTCG 60
QY 1640 TCGTTAGTTTGGG 1652
Db 61 TCGTTAGTTTGGG 73

RESULT 7
US-60-579-062-32930
; Sequence 32930, Application US/60579062
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Kovalic, David K.
; APPLICANT: Larosa, Thomas J.
; APPLICANT: Lu, Maolong
; APPLICANT: Munyikwa, Tichifa R. I.
; APPLICANT: Roberts, James K.
; APPLICANT: Wu, Wei
```

```

; APPLICANT: Zhang, Bei
; TITLE OF INVENTION: Methods for Genetic Control of Insect Infestations in Plants and
; FILE REFERENCE: Compositions thereof
; CURRENT APPLICATION NUMBER: US/60/579,062
; CURRENT FILING DATE: 2004-06-11
; NUMBER OF SEQ ID NOS: 41445
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 32930
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Diabrotica virgifera
US-60-579-062-32930

Query Match      3.3%; Score 55; DB 9; Length 533;
Best Local Similarity 47.8%; Pred. No. 0.0011;
Matches 160; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY 30 ACATGGTAAATTTTATAGTTAGTACTCTTGTAAATCATTTCAGAGGATGAAAAA 89
DB 54 AAAGAGATGATGTAAATAATAATAATAATAATAATAATAATAATAATAATA 113
QY 90 AATCGCAAGAAAGCAATATTTAAATGAATGCAATATACAAATTTAATFACAAAT 149
DB 114 ATTTTAAATATATATATATATATATATATATATATATATATATATATAT 173
QY 150 TATGTAAGATTCAATTTGTTAGTTTCATGAAATCAATTTCTAGAGTCATAATGCCT 209
DB 174 TTTTATTTTATAGAGATATATATATATATATATATATATATATATATATAT 233
QY 210 AAATTTAAATTTCTATCTTCTTTTCCCTTAAGAAAAAAGCAGCCATTAAGGACCA 269
DB 234 AAATTTATATATATAGAAATTTTAAATTAATAATAATAATAATAATAATAAT 293
QY 270 GAATGCGCGTGTCTCCATTTTATATATATAGATATAGTGTGTCTCCTAGGCCAC 329
DB 294 ATATTTGATTTATATATATATATATATATATATATATATATATATATATAT 353
QY 330 TCACCTCGCATGCTGTTCTTTTATACATGAAA 364
DB 354 TAAATATATATATATATATATATATATATATATATATATATATATATAT 388

RESULT 8
US-10-221-714B-498/c
; Sequence 498, Application US/10221714B
; GENERAL INFORMATION:
; APPLICANT: PIEPENBROCK, Alexander
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; FILE REFERENCE: tumor suppressor genes and oncogenes
; CURRENT APPLICATION NUMBER: US/10/221.714B
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 498
; LENGTH: 17280
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```

```

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714B-498

Query Match      3.1%; Score 50.6; DB 6; Length 17280;
Best Local Similarity 52.1%; Pred. No. 0.073;
Matches 113; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 38 TAAATTTTATAGTTAGTACTCTTGTAAATCATTTCAGAGGATGAAAAAATCGCAA 97
DB 5424 TAAATTAACAAATATAAATTTCACTCAAAAAAATAATAATAATAATAATAATA 5365
QY 98 GAAAGCAAAATATTTTAAATGAATGATGCAATATACAAATTTAATFACAAATATTAAG 157
DB 5364 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 5305
QY 158 ATTACATTTGTTAGTTTCATAGAAATCAATTTCTAGAGTCATAATATGCTTAATAA 217
DB 5304 AAAAAACATAAATCTCTTCACTTAATAATTTTACCATTCTTAAATATCTTAAAT 5245
QY 218 ATTCTATTCTATTTTCCCTTAAGAAAAAAGCAGCC 254
DB 5244 CAATTAATACATAAATAATAATAATAATAATAATAATACC 5208

RESULT 9
US-60-579-062-23067
; Sequence 23067, Application US/60579062
; GENERAL INFORMATION:
; APPLICANT: Baum, James A
; APPLICANT: Kovalic, David K
; APPLICANT: Larosa, Thomas J
; APPLICANT: Lu, Maolong
; APPLICANT: Munyikwa, Tichifa R. I.
; APPLICANT: Roberts, James K
; APPLICANT: Wu, Wei
; APPLICANT: Zhang, Bei
; TITLE OF INVENTION: Methods for Genetic Control of Insect Infestations in Plants and
; FILE REFERENCE: Compositions thereof
; CURRENT APPLICATION NUMBER: US/60/579,062
; CURRENT FILING DATE: 2004-06-11
; NUMBER OF SEQ ID NOS: 41445
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 23067
; LENGTH: 805
; TYPE: DNA
; ORGANISM: Diabrotica virgifera
US-60-579-062-23067

Query Match      3.0%; Score 50.4; DB 9; Length 805;
Best Local Similarity 46.6%; Pred. No. 0.02;
Matches 194; Conservative 0; Mismatches 221; Indels 1; Gaps 1;

QY 30 ACATGGTAAATTTTATAGTTAGTACTCTTGTAAATCATTTCAGAGGATGAAAAA 89
DB 202 AAAAAATCACAATCTTATGTAACAAATATTTTACATCATTTTTTTGTAATGATGC 261
QY 90 AATCGCAAGAAAGCAATATTTTAAATGAATGATGCAATATACAAATTTAATFACAAAT 149
DB 262 ATTCGCTTTTAAATATATATATATATATATATATATATATATATATATATATAT 321
QY 150 TATGTAAGATTCAATTTGTTAGTTTCATAGAAATCA-ATTTCTAGAGTCATAATGACC 208
DB 322 TATTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 381
QY 209 TAAATTAATAATCTATTTCTATTTTCTTAAGAAAAAAGCAGCCATTAAGGACCAT 268
DB 382 AATAAAAAATTTTATTAACATTTACTTTTAAATAAAAATAGGATCAACAATTTTCA 441
QY 269 AGAATGCGCGTGTCTCCATTTTATATATATAGAGATATAGTGTGTCTCCTAGGCCA 328
DB 442 TTAACATATGCTATCATTTATATATATATATATATATATATATATATATATAT 501

```



```

; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(207542)
; OTHER INFORMATION: n = A,T,C or G
US-10-902-387-493

Query Match      2.9%; Score 48.8; DB 6; Length 207542;
Best Local Similarity 49.2%; Pred. No. 0.65;
Matches 128; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 40 AATTTTATAGTTAGATTACTTCTTGTAATCAATTCAGAGGATGCAAAATTCGCAAGA 99
Db 168424 AGTAATTTTACATGATGATGCTAAATATATTAGCTTAAATTAAGAAATCTAAGTTAATA 168365

QY 100 AAGCAAAATATTTTAAATGAATGATGCAATATATACAAATTTTAAATTTACACAATTTATGTAAGAT 159
Db 168364 AAAACCTGACAATTTCTGGAATTTACACTGGATAAATTTTATGAAATTCGATATGAGATAT 168305

QY 160 TACATTTGTTTATGTTTCATAGAAATCAATTTCTAGAGTCATAATAATGCCTTAATTTAAAT 219
Db 168304 TGAATATTTAATTTGCAAAATTTTAAATTTTGCAGATAGAAAACAATAATGGTTTCTTCACATGACAT 168245

QY 220 TCTATTCTATTTTCTTAAAGAAAAAAGAGCCCATTAAGGACCATTAAGATGCGCGC 279
Db 168244 ACTATGCAATTTTAAATTTTGCAGATAGAAAACAATAATGGTTTCTTCACATGACCA 168185

QY 280 TGCTCCATTTTATATATAT 299
Db 168184 TGTAAATTTTCAATTCAT 168165

RESULT 14
US-10-893-315-134/c
; Sequence 134, Application US/10893315
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000786
; CURRENT APPLICATION NUMBER: US/10/893,315
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: 60/231,397
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 2172
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 134
; LENGTH: 207557
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(207557)
; OTHER INFORMATION: n = A,T,C or G
US-10-893-315-134

Query Match      2.9%; Score 48.8; DB 6; Length 207557;
Best Local Similarity 49.2%; Pred. No. 0.65;
Matches 128; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 40 AATTTTATAGTTAGATTACTTCTTGTAATCAATTCAGAGGATGCAAAATTCGCAAGA 99
Db 168424 AGTAATTTTACATGATGATGCTAAATATATTAGCTTAAATTAAGAAATCTAAGTTAATA 168365

QY 100 AAGCAAAATATTTTAAATGAATGATGCAATATATACAAATTTTAAATTTACACAATTTATGTAAGAT 159
Db 168364 AAAACCTGACAATTTCTGGAATTTACACTGGATAAATTTTATGAAATTCGATATGAGATAT 168305

QY 160 TACATTTGTTTATGTTTCATAGAAATCAATTTCTAGAGTCATAATAATGCCTTAATTTAAAT 219
Db 168304 TGAATATTTAATTTGCAAAATTTTAAATTTTGCAGATAGAAAACAATAATGGTTTCTTCACATGACAT 168245

QY 220 TCTATTCTATTTTCTTAAAGAAAAAAGAGCCCATTAAGGACCATTAAGATGCGCGC 279
Db 168244 ACTATGCAATTTTAAATTTTGCAGATAGAAAACAATAATGGTTTCTTCACATGACCA 168185

```

QY 280 TGCTCCATTTTATATAT 299
||| ||| ||| |||
Db 168184 TGTAAATTATTCATTCAT 168165

RESULT 15

US-10-902-387-465/c
; Sequence 465, Application US/10902387
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH CARDIOVASCULAR DISEASE AND LIPID DISORDERS, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL000787
; CURRENT APPLICATION NUMBER: US/10/902,387
; CURRENT FILING DATE: 2004-07-30
; PRIOR APPLICATION NUMBER: 60/231,399
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 6404
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 465
; LENGTH: 207557
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(207557)
; OTHER INFORMATION: n = A,T,C or G
US-10-902-387-465

Query Match 2.9%; Score 48.8; DB 6; Length 207557;
Best Local Similarity 49.2%; Pred. No. 0.65;
Matches 128; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
QY 40 AATTTTATAGTTAGATTACTTCTTGTATCATTTTCAGAGGATGAAAAAATCGCAAGA 99
Db 168424 AGTATTTTACATGATTAGATGCTAAATATTAGCTTAAAAATAGAAAACTAAGTTAATA 168365
QY 100 AAGCAAATATTTTAAATGAATGATGCAATATACAAATTTAATTACACAAATTATGTAAAGAT 159
Db 168364 AAACCTGACAAATTTGTGAAATTTACACTGGATAAATTTATTGAAATTTGATATGAGATAT 168305
QY 160 TACATTTGTTTATGTTTCATAGAAATCAATTTCTAGAGTCATATAATGCTAAATTTAAAT 219
Db 168304 TGAATATTTAATTTGCAAAATTTTAAATATATAGATACACATTTTATCTGCATATGACAT 168245
QY 220 TCTATTTCTATTTTTCCTTAAGAAAAAAGAGCCCATTTAAGGACCATTAGAATGCGGC 279
Db 168244 ACTATGCAATTTTAAATTTGAGATAAGAAAAACAATAATGTTTCTTCATGACCA 168185
QY 280 TGCTCCATTTTATATAT 299
||| ||| ||| |||
Db 168184 TGTAAATTATTCATTCAT 168165

Search completed: September 2, 2004, 12:48:44
Job time : 2374 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 07:29:38 ; Search time 3928 Seconds
(without alignments)
12604.760 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattcagggtcacataac.....ctcgtagttgggacggcg 1658

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estm:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_estl:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gssl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
C	1	751	45.3	779	28	CC418961	CC418961 PUEDX81TD
C	2	719.2	43.4	754	29	CG146202	CG146202 PUCV14TB
C	3	694.8	41.9	728	29	CG152000	CG152000 PUFT17TD
C	4	674	40.7	938	29	CG146205	CG146205 PUCV14TD

5	482.8	29.1	498	29	CG151999	CG151999 PUIFT17TB	
6	375.4	22.6	862	29	CG175932	CG175932 PUFY251TB	
C	7	373	949	28	BZ705097	BZ705097 PUCB143TD	
C	8	363.4	21.9	949	29	CG041829	CG041829 PUFW117TB
C	9	361.8	21.8	373	29	CG654144	CG654144 OGUDK20TV
C	10	358.8	21.6	879	29	CG118361	CG118361 PUFVU43TB
C	11	343.8	20.7	985	29	CG175933	CG175933 PUFY251TD
C	12	317.4	19.1	776	29	CG113123	CG113123 PUGY13TD
C	13	262.8	15.9	993	29	CC463184	CC463184 ZMMBBG034
C	14	250.4	15.1	844	29	CG096594	CG096594 PUFUD53TB
C	15	250.2	15.1	914	29	CG124360	CG124360 PUFUD11TB
C	16	212.4	12.8	728	29	CG164746	CG164746 PUUIU51TB
C	17	205.2	12.4	665	28	CC443358	CC443358 PUDIU72TD
C	18	194.4	11.7	775	28	BZ961663	BZ961663 PUBND60TD
C	19	194.4	11.7	826	28	BZ988278	BZ988278 PUBND60TD
C	20	192.2	11.6	732	28	CC352433	CC352433 PUXHD59TD
C	21	191.4	11.5	786	29	CG062076	CG062076 PUPDP96TD
C	22	191.4	11.5	858	29	CG062075	CG062075 PUPDP96TB
C	23	188	11.3	925	29	CG096596	CG096596 PUFUD53TD
C	24	175.8	10.6	884	28	CC006284	CC006284 PUDIN46TD
C	25	172.2	10.4	446	29	CG193725	CG193725 PUDJ37TD
C	26	170.8	10.3	438	29	CG193723	CG193723 PUDJ37TB
C	27	165	10.0	1032	29	CG046479	CG046479 PUFUV38TD
C	28	164.2	9.9	955	29	CG097258	CG097258 PUFWE94TB
C	29	138.4	8.3	656	29	CG050059	CG050059 PUFUA36TB
C	30	138.4	8.3	656	29	CG050060	CG050060 PUFUA36TD
C	31	127.6	7.7	866	29	CG118362	CG118362 PUFVU43TD
C	32	120.8	7.3	695	29	CC752611	CC752611 ZMMBB013
C	33	120.8	7.3	740	29	CG041830	CG041830 PUFW117TD
C	34	112	6.8	949	29	CG266345	CG266345 OG1A548TV
C	35	110.6	6.7	754	28	BZ828297	BZ828297 PUGAG16TD
C	36	110.6	6.7	807	28	BZ828291	BZ828291 PUGAG16TB
C	37	107.2	6.5	656	28	BH872036	BH872036 hm38d06.b
C	38	107	6.5	912	29	CG833955	CG833955 ZMMBB014
C	39	102.2	6.2	816	28	CC396173	CC396173 PUHLL28TD
C	40	101.6	6.1	528	10	AW216120	AW216120 687029C03
C	41	96.2	5.8	794	28	BZ991151	BZ991151 PUECF15TD
C	42	95.6	5.8	873	29	CG102130	CG102130 PUKCN52TD
C	43	95.4	5.8	523	28	BZ341248	BZ341248 ic44b07.g
C	44	92.2	5.6	870	11	AY106704	AY106704 Zea mays
C	45	89.4	5.4	871	28	CC366231	CC366231 PUHKK92TB

ALIGNMENTS

RESULT 1	CC418961	PUEDX81TD	ZM 0.6_1.0 KB	Zea mays	genomic clone ZMMBTA230M18,	779 bp	DNA	linear	GSS 19-MAY-2003
LOCUS	CC418961/c								
DEFINITION	CC418961				genomic survey sequence.				
ACCESSION	CC418961								
VERSION	CC418961.1				GI:30899051				
KEYWORDS	GSS.								
SOURCE	Zea mays								
ORGANISM	Zea mays								
REFERENCE									
AUTHORS									
TITLE									
JOURNAL									
COMMENT									

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES		Location/Qualifiers		Accession		Genomic survey sequence.	
source		1..779		CG146202		CG146202.1	
		/organism="Zea mays"		VERSION		GI:34036985	
		/mol_type="genomic DNA"		KEYWORDS		GSS.	
		/strain="B73"		SOURCE		Zea mays	
		/db_xref="taxon:4577"		ORGANISM		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.	
		/clone_lib="ZM061.0.KB"		REFERENCE		1 (bases 1 to 754)	
		/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high"		AUTHORS		Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J., and Bennett, J.	
		CoT selected genomic DNA library"		TITLE		Maize Genomics Consortium	
				JOURNAL		Unpublished (2003)	
				COMMENT		Other_GSSs: PUUCV14TD	
						Contact: Cathy Whitelaw	
						TIGR	
						9712 Medical Center Drive, Rockville, MD 20850, USA	
						Tel: 301-838-5843	
						Fax: 301-838-0208	
						Email: whitelaw@tigr.org	
						Seq primer: TR	
						Class: sheared ends.	


```

Db      274  CACGTAGTGTGCGCCCGAGGCTCTCGACACGGTGGCGCGCACGGGCGCGA 215
QY      1402  GGGACCTTGCAATTCGTCGGCTCGGTGCGACAGTCCCTGGTCCGGCTTGGCGGTTTCG 1461
Db      214  GGGACCTTGCAATTCGTCGGCTCGGTGCGACAGTCCCTGGTCCGGCTTGGCGGTTTCG 155
QY      1462  CCTGCGCTCGCGCGGAGCTGCGGTCTCCCGACCGAGCGCTATAAAAAGCGCGCGC 1521
Db      154  CTTGCGCTCGCGCGGAGCTGCGGTCTCCCGACCGAGCGCTATAAAAAGCGCGCGC 95
QY      1522  ACCTGATCTCCATCTCACCAAGCAAGCAGCAGCAACAGCAACCTAACTAAAGA 1581
Db      94  ACCTGATCTCCATCTCACCAAGCAAGCAGCAGCAACAGCAACCTAACTAAAGA 35
QY      1582  ACAGTAGTAGTCCCTGTGTAGTCTAGCAACGAT 1615
Db      34  ACAGTAGTAGTCCCTGTGTAGTCTAGCAACGAT 1

RESULT 3
CG152000/c
LOCUS      PU1F117D ZM.0.6.1.0 KB Zea mays genomic clone ZMMBTA0573C10,
DEFINITION genomic survey sequence.
ACCESSION  CG152000
VERSION     CG152000.1 GI:34042783
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (bases 1 to 728)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
            Maize Genomics Consortium
            Unpublished (2003)
            Other_GSSs: PU1F117B
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@cigr.org
            Seq primer: 1F
            Class: sheared ends.
            Location/Qualifiers
                1..728
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone.lib="ZM.0.6.1.0 KB"
                /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                Cot selected genomic DNA library"

ORIGIN
Query Match      41.9%; Score 694.8; DB 29; Length 728;
Best Local Similarity 99.0%; Pred. No. 5,5e-145;
Matches 721; Conservative 0; Mismatches 2; Indels 5; Gaps 2;

QY      601  AGCGACGAGTTGAAGCTACGGCGGTGGCTGCTGCGCGGTGCTAGCAGCGGAC 660
Db      728  AGCGACGAGTTGAAGCTACGGCGGTGGCTGCTGCGCGGTGCTAGCAGCGGAC 669
QY      661  GACGTCGACGCTGAGCTGCTGCTACTTTGTTTGTCTCTGCTGCTGCGATG 720
Db      668  GACGTCGACGCTGAGCTGCTACTTTGTTTGTCTCTGCTGCTGCGATG 609
QY      721  ATGCAGTAAATAAACGTTGTATCTCATGTGATTGATCGAGCTGTGTCGGCGATTTCAA 780

```

```

Db      608  ATGCGAGTAAATAAACGTTGTATCTCATGTGATTGATCGAGCTGTGTCGGCGATTTCAA 549
QY      781  ATCAGTAAACAATCACATGTGCAATCCCATCGACTTCGTGTCGGCGATAAATCTGTAGGG 840
Db      548  ATCAGTAAACAATCACATGTGCAATCCCATCGACTTCGTGTCGGCGATAAATCTGTAGGG 489
QY      841  ATTGTGCTTTTCATGTATGTCTGTAATAGAGAGAAAGAATAATATTACATAAATAT 900
Db      488  ATTGTGCTTTTCATGTATGTCTGTAATAGAGAGAAAGAATAATATTACATAAATAT 429
QY      901  GAGTTCAACATCTTCTTAAATACCTTTGCTCGATCTCTTCTAGTAGAGCCAGTGGCTGA 960
Db      428  GAGTTCAACATCTTCTTAAATACCTTTGCTCGATCTCTTCTAGTAGAGCCAGTGGCTGA 369
QY      961  TAACGTGTTAAGAACCCCTTTGTTACCGAGTGTAGTCCAA-GGCTGCTTAACTGAATAATA 1019
Db      368  TAACGTGTTAAGAACCCCTTTGTTACCGAGTGTAGTCCAAAGGGCTGTTAACTGAATAATA 309
QY      1020  GTAGAACACAGGACAGAGAGTGTAGAGAGAACTGATCTTTGTTGTTACTATATGTTGCT 1079
Db      308  GTAGAACACAGGACAGAGAGTGTAGAGAGAACTGATCTTTGTTGTTACTATATGTTGCT 249
QY      1080  GCTCTCCAAAGCTTACATGATATGGGATCTCTCTCTATTTATAGACAAACTAGGGTT 1139
Db      248  GCTCTCCAAAGCTTACATGATATGGGATCTCTCTCTATTTATAGACAAACTAGGGTT 189
QY      1140  TCAGGCATATGGGCCACATAGGCTTCTCTGGCCCAAGAAAGGTTTCTTAACTACTACCATC 1199
Db      188  TCAGGCATATGGGCCACATAGGCTTCTCTGGCCCAAGAAAGGTTTCTTAACTACTACCATC 129
QY      1200  TTCAGAGGCGGTGCTGCTGCTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1255
Db      128  TTCAGAGGCGGTGCTGCTGCTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 69
QY      1256  ATACGGGTCTACGCTGCACGCGCAGCATGTGTGCGCGAGGGGCGATCTCTCCCGGGGA 1315
Db      68  ATACGGGTCTACGCTGCACGCGCAGCATGTGTGCGCGAGGGGCGATCTCTCCCGGGGA 9
QY      1316  ATGCGCGC 1323
Db      8  ATGCGCGC 1

CG146205      938 bp DNA linear GSS 21-AUG-2003
PUUCV14TD ZM.0.6.1.0 KB Zea mays genomic clone ZMMBTA0639C04,
DEFINITION genomic survey sequence.
ACCESSION  CG146205
VERSION     CG146205.1 GI:34036988
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 938)
            Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
            Maize Genomics Consortium
            Unpublished (2003)
            Other_GSSs: PUJCV14TB
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@cigr.org
            Seq primer: 1F
            Class: sheared ends.
            Location/Qualifiers
                1..938
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone.lib="ZM.0.6.1.0 KB"
                /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                Cot selected genomic DNA library"

ORIGIN
Query Match      41.9%; Score 694.8; DB 29; Length 728;
Best Local Similarity 99.0%; Pred. No. 5,5e-145;
Matches 721; Conservative 0; Mismatches 2; Indels 5; Gaps 2;

QY      601  AGCGACGAGTTGAAGCTACGGCGGTGGCTGCTGCGCGGTGCTAGCAGCGGAC 660
Db      728  AGCGACGAGTTGAAGCTACGGCGGTGGCTGCTGCGCGGTGCTAGCAGCGGAC 669
QY      661  GACGTCGACGCTGAGCTGCTGCTACTTTGTTTGTCTCTGCTGCTGCGATG 720
Db      668  GACGTCGACGCTGAGCTGCTACTTTGTTTGTCTCTGCTGCTGCGATG 609
QY      721  ATGCAGTAAATAAACGTTGTATCTCATGTGATTGATCGAGCTGTGTCGGCGATTTCAA 780

```

source		1. .938		Zea mays		ORGANISM	
		/organism="Zea mays"		Bukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
		/mol_type="genomic DNA"		Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD			
		/strain="B73"		clade; Panicoidae; Andropogoneae; Zea.			
		/db_xref="taxon:4577"		1 (bases 1 to 498)			
		/clone="ZMMBta0639C04"		Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,			
		/clone_lib="ZM 0.6 1.0 KB"		Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J., and			
		/notes="vector: PCR4-TOFO; Site_1: EcoRI; 0.6-1.0 kb high		Bennetzen, J.			
		Cor selected genomic DNA library"		Maize Genomics Consortium			
				Unpublished (2003)			
				Other GSSs: PU1F117TD			
				Contact: Cathy Whitelaw			
				TIGR			
				9712 Medical Center Drive, Rockville, MD 20850, USA			
				Tel: 301-838-5843			
				Fax: 301-838-0208			
				Email: whitelaw@tigr.org			
				Seq primer: TR			
				Class: sheared ends.			
				Location/Qualifiers			
				1. .498			
				/organism="Zea mays"			
				/mol_type="genomic DNA"			
				/strain="B73"			
				/db_xref="taxon:4577"			
				/clone="ZMMBta0573C10"			
				/clone_lib="ZM 0.6 1.0 KB"			
				/notes="vector: PCR4-TOFO; Site_1: EcoRI; 0.6-1.0 kb high			
				Cor selected genomic DNA library"			
				Query Match		29.1%; Score 482.8; DB 29; Length 498;	
				Best Local Similarity		99.4%; Pred. No. 1.8e-97;	
				Matches 495; Conservative		0; Mismatches 2; Indels 1; Gaps 1;	
				ORIGIN			
				QY		266 ATTAGAATCGCGCTGCTCCA-TTTTATATATATAGATATAGTGTGCTCACTAG 324	
				Db		1 ATTAGAATCGCGCTGCTCCA-TTTTATATATATAGATATAGTGTGCTCACTAG 60	
				QY		325 GCCACTCCACCTCGCATGGTGTCTTTATCATGAAAAAATGAGCGCTGCTTTGCA 384	
				Db		61 GCCACTCCACCTCGCATGGTGTCTTTATCATGAAAAAATGAGCGCTGCTTTGCA 120	
				QY		385 TTAAGGCTCTAGAGATGTACAGTGCACACGCTTCGTATTCACTGCACGAGTAATTTTT 444	
				Db		121 TTAAGGCTCTAGAGATGTACAGTGCACACGCTTCGTATTCACTGCACGAGTAATTTTT 180	
				QY		445 TTAATGCACGAGGCGCTTTGTATTCTCTAGCAGCTCTGCGCGCCGCTGCTGTGTACG 504	
				Db		181 TTAATGCACGAGGCGCTTTGTATTCTCTAGCAGCTCTGCGCGCCGCTGCTGTGTACG 240	
				QY		505 GCAGTCAGGAGGCGAGCAGCTCTAGTGCACACGCTGCTGCTGCCACCGCTGA 564	
				Db		241 GCAGTCAGGAGGCGAGCAGCTCTAGTGCACACGCTGCTGCTGCCACCGCTGA 300	
				QY		565 TGGCAAGGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 624	
				Db		301 TGGCAAGGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360	
				QY		625 GTGGCTGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 684	
				Db		361 GTGGCTGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420	
				QY		685 GCTACTTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744	
				Db		421 GCTACTTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480	
				QY		745 CTCATGTGATTGATCGAC 762	
				Db		481 CTCATGTGATTGATCGAC 498	
				RESULT 6			

source

1. .938
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBta0639C04"
/clone_lib="ZM 0.6 1.0 KB"
/notes="vector: PCR4-TOFO; site_1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

ORIGIN

Query Match

40.7%; Score 674; DB 29; Length 938;

Best Local Similarity 98.3%; Pred. No. 2.5e-140;

Matches 702; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY

1 GAATTCACGGCTCACAAATACCAATCATCTACATGTGATAAATTTTATAGTTAGTACTT 60

Db

225 GAATTCACGGCTCACAAATACCAATCATCTACATGTGATAAATTTTATAGTTAGTACTT 284

QY

61 CTTGTAATCATTTTCAGAGGATGAATAAATAATCGCAAGCAAAATATTTTAAATGAAT 120

Db

285 CTTGTAATCATTTTCAGAGGATGAATAAATAATCGCAAGCAAAATATTTTAAATGAAT 344

QY

121 GATGCAATATACAAATTTAATATACAAATATATGTAGATTTACATTTGTTAGTTTCATAGA 180

Db

345 GATGCAATATACAAATTTAATATACAAATATATGTAGATTTACATTTGTTAGTTTCATAGA 404

QY

181 AATCAATTTCTAGAGTCATTAATATGCTAAATTAATAATTTCTATTTCTTTTCTTAAAG 240

Db

405 AATCAATTTCTAGAGTCATTAATATGCTAAATTAATAATTTCTATTTCTTTTCTTAAAG 464

QY

241 AAAAAAAGACCCATTAAAGGGACCATTAAGATCGCGCTGCTCCA-TTTTATATATAT 299

Db

465 AAAAAAAGACCCATTAAAGGGACCATTAAGATCGCGCTGCTCCA-TTTTATATATAT 524

QY

300 AGAGATATGATGTTGTTGCTCACTAGGCCACTCCACCTTGCATGCGTGTCTTTATTACAT 359

Db

525 AGAGATATGATGTTGTTGCTCACTAGGCCACTCCACCTTGCATGCGTGTCTTTATTACAT 584

QY

360 GAAAAAATGACGCGCTGCTTTTGCATTAAGGCTTAGAGATGTACGAGTGCACGCTTCG 419

Db

585 GAAAAAATGACGCGCTGCTTTTGCATTAAGGCGCTTAGAGATGTACGAGTGCACGCTTCG 644

QY

420 TATTCACTGCACGAGTAATTTTTTTTAAATGCACGAGGCGCTTTGTATTCCTCAGCAGC 479

Db

645 TATTCACTGCACGAGTAATTTTTTTTAAATGCACGAGGCGCTTTGTATTCCTCAGCAGC 704

QY

480 TCTCGCGCCGCTGCTTCGTGACGGCAGTACGAGGCGACGACGAGCTCTACGTGCCAC 539

Db

705 TCTCGCGCCGCTGCTTCGTGACGGCAGTACGAGGCGACGAGCAGCTCTACGTGCCAC 764

QY

540 GTGCTGTGCTGCTGCCACACCGTGTATGTCAGAGGCTTGCTTGCTGCGTGCCTGCGTGG 599

Db

765 GTGCTGTGCTGCTGCCACACCGTGTATGTCAGAGGCTTGCTTGCTGCGTGCCTGCGTGG 824

QY

600 CAGCGACGAGGTTGAAGCTACGGCGGTGGCTTGCTGCTGCGCGGTGCTAGCAGCGGCA 659

Db

825 CAGCGACGAGGTTGAAGCTACGGCGGTGGCTTGCTGCTGCGCGGTGCTAGCAGCGGCA 884

QY

660 CGACGTCCAGCTGATGCTAGGTTGCTAC-TTTTGTTTTCTTTTCTCTGTGTC 712

Db

885 CGACGTCCAGCTGATGCTAGGTTGCTACTTTTGTGTTTTTGTCTCTGTGTC 938

RESULT 5

CG151999

LOCUS

CG151999
PU1F117B_ZM_0.6_1.0_KB_Zea mays genomic clone ZMMBta0573C10, genomic survey sequence.
CG151999.1
GI:34042782
GSS.
Zea mays

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

CG175932	CG175932	862 bp	DNA	linear	GSS 21-AUG-2003	
LOCUS	PUFY251TB.ZM.0.6.1.0.KB.Zea.mays.genomic.clone.ZMMBTa0765I06,					
DEFINITION	genomic survey sequence.					
ACCESSION	CG175932					
VERSION	CG175932.1	GI:34066730				
KEYWORDS	GSS.					
SOURCE	Zea mays					
ORGANISM	Zea mays					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.					
REFERENCE	1 (bases 1 to 862)					
AUTHORS	Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.					
TITLE	Maize Genomics Consortium					
JOURNAL	Unpublished (2003)					
COMMENT	Other_GSSs: PUFY251TD Contact: Cathy Whitelaw TIGR					
	9712 Medical Center Drive, Rockville, MD 20850, USA					
	Tel: 301-838-5843					
	Fax: 301-838-0208					
	Email: whitelaw@cigr.org					
	Seq primer: TR					
	Class: Sheared ends.					
FEATURES	Location/Qualifiers					
source	1..862					
	/organism="Zea mays"					
	/mol_type="genomic DNA"					
	/strain="B73"					
	/db_xref="taxon:4577"					
	/clone="ZMMBTa0765I06"					
	/clone_lib="ZM.0.6.1.0.KB"					
	/notes="Vector: pCR4-TORO; Site_1: EcoRI; 0.6-1.0 kb high					
	Cor selected genomic DNA library"					
ORIGIN						
	Query Match	22.6%;	Score 375.4;	DB 29;	Length 862;	
	Best Local Similarity	82.4%;	Pred.No. 2.1e-73;			
	Matches 520;	Conservative	0;	Mismatches 91;	Indels 20;	Gaps 7;
QY	576	GCTTGCTTCGTCCTGCGTGTGCGACGACGAGGTTGAAGCTACGCGGTGCTGCTTG	635			
Db	85	GCTCGCTTGACGCTGCTGTGCGGTAGCAGCGAGGCCGGGC-ACGACAGTAGGTACTTG	143			
QY	636	CCTGCGCGCTGTAGCAGCGACGACGCTGCGCTGATGCTGAGCTTGCTACTTTTGT	695			
Db	144	CCTGTGTGCGTGAAGCGTGGCACTGAAGCAGCCCT-----ATAGTTTTTT	190			
QY	696	TTTGTGTTTTCTGTGCTTCGGATGATGTCAGTAAATAAATACGTTGCTATCTCATGTGATT	755			
Db	191	CTTCTTTTCTATGTTCTTCAGATGTCAGTGAATAAATGTGTTCATGTGATTGA	250			
QY	756	GATCGACGTGTGCGGGAATTCAAATCAGTAAACAATCAATGTCATGCCCATCCGACTT	815			
Db	251	TCGACGTGTGTGTCGGCGAATTCAAATCAGCAACATCACGTGTGCATCCCATCGACTT	310			
QY	816	CGTGTGCGCGATAAATCTGCT-AGGGAATTTGTGCTTTTTCATGTATGTGCTTGAATAGAGA	874			
Db	311	CGTGTACGCAATAAATCTACTCGGGGATTTGTGCTTTTTCATGTATGTGCTTGAATAGAGA	370			
QY	875	GGAAAGAAATAATTA--CATAAAATATGAGTTCAACATCTTTCTTAATACCTTTTGCTC	932			
Db	371	GAAGAAGCAATATTCGCAATAAATATGAGTTCAACGCTTTCTTAATACCTTTTGCTC	430			
QY	933	GATCTCCTTAGTAGAGCCAGTCGCTGATTAACGTGTTAAGAACCCCTTGTATCCGAGTGT	992			
Db	431	AATCTCCTTAGTAGAGCCAGTCGCTGATTAAGTGTTAAGAACCCCTTGTATCCGAGTAT	490			
QY	993	AGTCCAAAGG-CTGCTTAACCTGAATAATAGTAGAACACAGGAGACAGAGTGTAGAGAGG	1051			

```

QY 241 AAAAAAAAAAGCCATTAAAGGACCAATTAGATGCGCGTGTCTCCA-TTTTTTATATAT 299
Db 169 AAAAAAAAAAGCCCAATTAGGACCAATTAGATGCGCGTGTCTCCAATTTTTTATATAT 110
QY 300 AGAGATATGAGTGTGTGCTCACTAGGCCCACTCCACCTCGCATGCGTGTCTTTTATTACAT 359
Db 109 AGAGATATGAGTGTGTGCTCACTAGGCCCACTCCACCTTCGATGCGTGTCTTTATTACAT 50
QY 360 GAAAAAATGAGCGCGTGTCTTTTGCATTAAGGCTTAGAGAT 400
Db 49 GAAAAAATGAGCGCGTGTCTTTTGCATTAAGGCTTAGAGAT 9

RESULT 8
CG041829/c
LOCUS
DEFINITION
  CG041829 949 bp DNA linear GSS 19-AUG-2003
  PUFW117B ZM 0.61.0 KB Zea mays genomic clone ZMBMa0748C09,
  genomic survey sequence.
ACCESSION
  CG041829
VERSION
  CG041829.1 GI:33913985
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
  1 (bases 1 to 949)
  Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
  Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
  Bennetzen, J.
  Maize Genomics Consortium
  Unpublished (2003)
  Other_GSSs: PUFW117D
  Contact: Cathy Whitelaw
  TIGR
    9712 Medical Center Drive, Rockville, MD 20850, USA
    Tel: 301-838-5843
    Fax: 301-838-0208
    Email: whitelaw@tigr.org
    Seq primer: TR
    Class: sheared ends.
  FEATURES
    source
      1..949
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /strain="B73"
        /db_xref="taxon:4577"
        /clone="ZMBMa0748C09"
        /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
        Cot selected genomic DNA library"
  ORIGIN
    Query Match 21.9%; Score 363.4; DB 29; Length 949;
    Best Local Similarity 84.5%; Pred. No. 1e-70;
    Matches 432; Conservative 0; Mismatches 76; Indels 3; Gaps 2;

QY 700 GTTTCTCTGTGCTGCGATGTCAGTAAATAAAGCTGTGTATCTCATGTGATTC 759
Db 632 GTTGTGTGTGCTCTTTTGGTTATATGATAAAATAAAGCTGTGTCATATGTGATTGATC 573
QY 760 GAGTGTGTGCGGATTTCAAATCAGTAACAATCACATGTGCATCCCACTCGACTTCGTG 819
Db 572 GAGTGTGTGAGGACTTCAAGTCAGCAGCAATCACATGTACACTCCATCGACTTCGTG 513
QY 820 TCGCGGATAAATCTGTAGGATTTGTCTTTTATGTATGTCTTGAATAGAGAGAAA 879
Db 512 CCGGCGATAAATCAGCTTGGGATGTGATTTCTCGTGCATGTGCTGAATAAGAAATAAG 453
QY 880 GAATATATTACAT--AAATATGAGTTCAACATCTTTCTTAATACCTTTTGTCTCATCT 937
Db 452 AACACACTATACATAGATAATATAGTTTCAGCACCTTTCTTAGTACCTTTTGTCTCATCT 393

```

```

QY 938 CCTTACTAGAGCACGTGCGTGATTAACGTTTAAAGAACCCCTTGTGTACCGAGTGTAGTCC 997
Db 392 TCTTAGTAGAGCCACGTGTGTGTGATTAACGTTTAAAGAACCCCTTGTGTACCGAGTGTAGTCC 333
QY 998 AA-GGCTGCTTAACTGAATAATAGTAGAACAACAAGAGACAAGAGTGTAGAGAGAACTG 1056
Db 332 AAGGGCTGCTTAACTGAATAATAGTAGAACAACAAGAGACAAGAGTGTAGAGAGAACTG 273
QY 1057 ATCTTTGTACTATATGTGTCTCTCTCAAGGTTACATGATATGGGATCTCTCTC 1116
Db 272 ATCTTTGTACTATATGTGTCTCTCTCAAGGTTACATGATATGGGATCTCTCTC 213
QY 1117 TATTATAGACAAAACCTAGGCTTTTACAGCATATGGGCCACATAGGCTTCTCTGGCCCAAG 1176
Db 212 TATTATAGACAAAACCTAGGCTTTTACAGCATATGGGCCACATAGGCTTCTCTGGCCCAAG 153
QY 1177 AAAGGTTTCTTAACACTACCATCTTCCAGGC 1207
Db 152 AAAGGTTTCTTAACAGAAAACAACATGTAGCC 122

RESULT 9
CC654144
LOCUS
DEFINITION
  CC654144 373 bp DNA linear GSS 19-JUN-2003
  OGUDK20TV ZM 0.71.5 KB Zea mays genomic clone ZMBMa0415C15,
  genomic survey sequence.
ACCESSION
  CC654144
VERSION
  CC654144.1 GI:32057168
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
  1 (bases 1 to 373)
  Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
  Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
  Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
  Consortium for Maize Genomics
  Unpublished (2002)
  Contact: Cathy Whitelaw
  TIGR
    9712 Medical Center Drive, Rockville, MD 20850, USA
    Tel: 301-838-5843
    Fax: 301-838-0208
    Email: whitelaw@tigr.org
    Seq primer: TF
    Class: sheared ends.
  FEATURES
    source
      1..373
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /strain="B73"
        /db_xref="taxon:4577"
        /clone="ZMBMa0415C15"
        /clone_lib="ZM 0.71.5 KB"
        /note="Vector: pBCSk-; Site 1: HincII; 0.7-1.5 kb
        methylation filtered genomic DNA library"
  ORIGIN
    Query Match 21.8%; Score 361.8; DB 29; Length 373;
    Best Local Similarity 99.5%; Pred. No. 2.5e-70;
    Matches 363; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1288 TCGCGCAGGGGATCTCTGCGGGGAATGCGCAGGCGCATCGTCGACACGACGTA 1347
Db 1 TCGCGCAGGGGATCTCTGCGGGGAATGCGCAGGCGCATCGTCGACACGACGTA 60
QY 1348 CGTGTGCGCCCGCGAGGCTCTCTCGCACACGCGTTCGCGCGCACCGGGCCCGAGGACC 1407
Db 61 CGTGTGCGCCCGCGAGGCTCTCTCGCACACGCGTTCGCGCGCACCGGGCCCGAGGACC 120

```


Qy	778	CAAAATCAGTAAACAATCAATGTGCATCCCATCGACTTCGTGCGGAGTAAATCTGCCT-	8336
Db	922	CAAAATCAGCAACAATCAACGTGTGCATCCCATCGACTTCGTGAGCAATAAATCTACTC	863
Qy	837	AGGGATTGTGCTTTTCATGTATGTGCTTGAATAGAGAGGAAAGAAATAATATTA--CATA	894
Db	862	GGGGATTGTGCTTTTCATGTATGTGCTTGAATAGAGAGAAAGAACAAATTTACGGATA	803
Qy	895	AAATATGAGTTCAACATCTTTTCTTAATACCTTTTGTCTCGATCTCCTTAGTAGGCCACGT	954
Db	802	AAATATGAGTTCAACGCTCTTCTTAATACCTTTTGTCTCAATCTCCTTAGTAGGCCACGT	743
Qy	955	GGCTGATAACGCTGTATGAACCCCTTGTACCGAGTGTAGTCCAAAG-CTGCTTAACTGA	1013
Db	742	GGCTGATAAAGTGTATGAACCCCTTGTACCGAGTATAGTCCAAAGACTCTTTAACTGA	683
Qy	1014	ATAATAGTAGAAACAACAAGGAGACAAGAGTGTAGAGAGAACTGATCTTTTGTACTATAT	1073
Db	682	ATAATAGTAGAAACAACAAGGAGACAAGAGTGTAGAGAGGACTGATCTTTTAACTGTAT	623
Qy	1074	GGTGCTCCTCTCCAAAGGTTACATGATATGGGAGATCTCCTCTCTATTTTATAGAC-AAAAC	1132
Db	622	GGTGCTCTTGTGCC--CAGGTTTACATGATATGGGGATCTCCTCTCTATTTATAGACAAAAC	564
Qy	1133	TAGGGTTTCAGGCATATGGGCCACATAGGCCCTTCCCTGGCCCAAGAAAGGTTTCTTAAACAC	1192
Db	563	TAAGGTTTCGATATATGGGCCACATGGGCCCTTCCCTGGCCCAATAAAGGTTTCTTAAACAC	504
Qy	1193	TACCATCTT 1201	
Db	503	GCCCAAGCT 495	

RESULT 12			
CG113123/c			
LOCUS	CG113123	776 bp	DNA linear GSS 20-AUG-2003
DEFINITION	PUGV13TD_ZM_0.5_1.0_KB_Zea mays genomic clone ZMMBTa0666C01,		
	genomic survey sequence.		

CG113123
CG113123.1 GI:33996560
VERSION GSS
KEYWORDS Zea mays
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE
AUTHORS Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 776)
Whitellaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.

TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
CONTACT Cathy Whitellaw
TGR

9/12 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitellaw@tigr.org
Seq primer: TF
Class: sheared ends.

```

FEATURES
    source
        Location/Qualifiers
            1..776
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZMMBT30666C01"
                /clone_lib="ZM_0.6-1.0_KB"
                /note=vector: pCR4-TOPO; site 1: EcoRI; 0.6-1.0 kb high
                Cor selected genomic DNA library"

```

Query Match

19.1%; Score 317.4; DB 29; Length 776;

[illegible][illegible]

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1. (bases 1 to 993)
Routard, A.K., Young, S., Kavchok, S., Keizer, G., Bronzino, A.C.,
Bharti, A.K., Fuks, G., Yu, Y., Wing, R., and Messing, J.
Sequencing of the maize genome at PGIR (2003b)
Unpublished (2003)
Contact: Bharti, A.K.
Dr. Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA

```

FEATURES
source
1. .993
/organism="Zea mays subsp. mays"
/mol_type="genomic DNA"
/cultivar="B73"
/sub_species="mays"
/db_xref="taxon:4578"
/clone="ZMMBBC0349F01"

```

```
/lab_host="E. coli DH10B"
/clone_lib="ZMBBc"
/note="Vector: pTARAC1.3; Site_1: BamHI; Site_2: BamHI"

ORIGIN
Query Match 15.9%; Score 262.8; DB 29; Length 993;
Best Local Similarity 73.4%; Pred. No. 3.7e-48;
Matches 496; Conservative 0; Mismatches 142; Indels 38; Gaps 11;

QY 532 GTCCACACGCGCTGGCTGCTGCCACCGCTGATGCGAAGCTTGCTTGGTCCGTCGCT 591
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 592 GCGTGTGGCAGCAGCGAGGTGAAGCTACGGCGGTGCTTGGCTGCGCGGTGCTAG 651
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 671 TCGTAGAGCAGCGGTGAGCGCACTGCTACGACGGCTGCTTGGCTTCCATCGCTAG 612
QY 652 CAGCGGACAGCTGCGACGCTGATGCTAGCTTGGCTTGGCTTGGCTTGGCTTGGCT 711
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 611 TACAGCGACT-----ACGCTGACGCTAGCCTTACTACTGTTGTTT----- 565
QY 712 CTGCGATGATGAGTAATAAATCAATCAATGTCATCCCATCGACTTCTGTCGCGATTA 829
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 564 -----AATGATGAGAAATAAATCAATGTCATCCCATCGACTTCTGTCGCGATTA 829
QY 770 GCGATTTCAATCAATCAATCAATGTCATCCCATCGACTTCTGTCGCGATTA 829
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 509 GCGATTTAAAGTCAGCGCAATCATGCTGCA-CCCATCGACTTCTGTCGCGATTA 451
QY 830 ATCTGCTAGGATTTGCTTTTCAATGTCATCCCATCGACTTCTGTCGCGATTA 451
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 450 AATCTGCTAGGATTTGCTTTTCAATGTCATCCCATCGACTTCTGTCGCGATTA 451
QY 885 ATATTACATAAATATGATGTCATCAATCTTCTTAATACCTTCTGTCGATCTCTTACT 944
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 390 AGATCATATAATATGATGTCATCAATCTTCTTAATACCTTCTGTCGATCTCTTACT 944
QY 945 AGAGCCAGCTGG-----TGATAACGCTGTTAAGAACCCCTTGTACCAGGTAGTCC 997
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 998 AA-GGCTGCTTAATGATTAATAGTAGAACAAGAGAGCAAGAGTGTAGAGGAACTG 1056
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 271 AAGGGCTCCTTAATGATTAATAGTAGAACAAGAGAGCAAGAGTGTAGAGGAACTG 212
QY 1057 ATCTTTGTTACTATATGTCGCTCTCCAAAGTGTACATGATATGGGATCTCTCTC 1116
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 211 ATCTTTTACTATGATGTCGCTCTCCAAAGTGTACATGATATGGGATCTCTCTC 1116
QY 1117 TATTTATAGACAA-AACTAGGTTTTCAGGCATATGGCCACATAGGCTTCTGCGCCAA 1175
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 153 TATTTATAGACAA-AACTAGGTTTTCAGGCATATGGCCACATAGGCTTCTGCGCCAA 1175
QY 1176 GAAAGTTTCTTAACA 1191
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 93 GACAAGATTTCTTAA 78

RESULT 14
CG096594
LOCUS
DEFINITION
PUFUD53TB ZM 0.6 1.0 KB Zea mays genomic clone ZMBB0733J10,
genomic survey sequence.
ACCESSION
CG096594
VERSION
CG096594.1 GI:3397888
KEYWORDS
GSS.
SOURCE
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 844)
Whitehouse, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,

Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUFUD53TD
Contact: Cathy Whitelaw
TIGR Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..844
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMBB0733J10"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN
Query Match 15.1%; Score 250.4; DB 29; Length 844;
Best Local Similarity 77.8%; Pred. No. 2.2e-45;
Matches 343; Conservative 0; Mismatches 88; Indels 10; Gaps 3;

QY 768 TCGGGATTTCAATCAATCAATGTCATCCCATCGACTTCTGTCGCGAT 827
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 203 TTGGGACTTCAATGTCAGCAACACACGCTTCTTATCGACTTCCGAGTCGCGAT 262
QY 828 AATCTGCTAGGATTTGCTTTTCAATGTCATCCCATCGACTTCTGTCGCGAT 827
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 263 AATCTGCTAGGATTTGCTTTTCAATGTCATCCCATCGACTTCTGTCGCGAT 827
QY 886 -----TATTAATCAATCAATGTCATCCCATCGACTTCTGTCGCGAT 827
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 323 ATCATGTAATAGAGACATAGTTCAAAAGATTTCTTCAATACCTTTGTTGATCTCT 382
QY 940 TTAGTAGAGCAGCGCTGATAGCTGTTAAGAACCCCTTGTACCAGGTAGTCCAA 999
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 393 TGAATAGAGTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
QY 1000 -GGCTGCTTAACTGAATAGTAGAACAAGAGAGCAAGAGTGTAGAGGAACTGAT 1058
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 442 GGGCTGCTTAACTGAATAGTAGAACAAGAGAGCAAGAGTGTAGAGGAACTGAT 501
QY 1059 TCTTTGTTACTATATGTCGCTCTCCAAAGTGTACATGATATGGGATCTCTCTCTA 1118
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 502 TCTTTGTTACTATATGTCGCTCTCCAAAGTGTACATGATATGGGATCTCTCTCTA 561
QY 1119 TTTATAGACAAACTAGGTTTTCAGGCATATGGCCACATAGGCTTCTGCGCCAA 1178
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 562 TTTATAGACAAACTAGGTTTTCAGGCATATGGCCACATAGGCTTCTGCGCCAA 1178
QY 1179 AGGTTTCTTAACACTTACCATC 1199
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 622 AGGTTTCTTAACACTTACCATC 642

RESULT 15
CG124360/c
LOCUS
DEFINITION
CG124360 ZM 0.6 1.0 KB Zea mays genomic clone ZMBB0654A22,
genomic survey sequence.
ACCESSION
CG124360
VERSION
CG124360.1 GI:34007797
KEYWORDS
GSS.
SOURCE
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 914)
REFERENCE
AUTHORS
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.
TITLE
JOURNAL
COMMENT
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUUFD11TD
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: Sheared ends.
FEATURES
source
1..914
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa0654A22"
/clone_lib="ZM 0.6-1.0 KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"
ORIGIN
Query Match 15.1%; Score 250.2; DB 29; Length 914;
Best Local Similarity 77.8%; Pred. No. 2.4e-45;
Matches 343; Conservative 0; Mismatches 88; Indels 10; Gaps 3;
QY 768 TCGCGGATTTCAATCAGTAAACATGATGTCATCCCATCGACTTCGTGCGGGAT 827
Db 533 TTGGCGACTTCATCTCAGCAACACTCAGGTGTTCTATTTCTGACTTCGAGTCGGCGAT 474
QY 828 AATCTCTAGGGATTTGCTTTTCAATGATGCTTGATGCTTGATAGAGGAAGAAATAA-- 885
Db 473 ATAATCAGCTAAGAAATGATCATATTTGTAATATGAAAGGAAAAACAGGATTCAAAAT 414
QY 886 -----TATTACATAAAATATGAGTTCACATCTTTCTTAATACCTTTTGTCTGATCTCC 939
Db 413 ATCATGTATATAGAGAACTAGTTCAAGATTCTTCTCATATACCTTTGTTGATCTTCT 354
QY 940 TTATAGAGCCACGTGCGTGATACAGTGTAAAGAACCCCTTTTACCGAGTGTAGTCCAA 999
Db 353 TGAATAGAGTCTCATGC-TGATAACGTGTTAAGAACCCCTTGTACCGAGTGTAGTCCAA 295
QY 1000 -GGCTGCTTAAGTGAATATAGTAGAACAACAGGAGACAGAGTGTAGAGAGGAAGTAT 1058
Db 294 GGGCTGCTTAAGTGAATATAGTAGAACAACAGGAGACAGAGTGTAGAGAGGAAGTAT 235
QY 1059 TCCTTTGTACTATATGCTGCTCTCCAAAGGTTACATGATATGGGATCTCCTCTCTA 1118
Db 234 TCCTTTGTACTATATGCTGCTCTCTCCAAAGGTTACATGATATGGGATCTCCTCTCTA 175
QY 1119 TTTATAGACAAAATAGGGTTTCAGGCATATGGCCACATAGGCTTCTCGGCCCAAGAA 1178
Db 174 TTTATAGACAAAATAGGGTTTCAGGCATATGGCCACATAGGCTTCTCGGCCCAAGAA 115
QY 1179 AGGTTTCTTAACACTACATC 1199
Db 114 AGGTTTCTTAACATAGCAAC 94

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 12:09:10 ; Search time 19145 Seconds
(without alignments)
3753.604 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattcaggtcacaatac.....ctcgttagtttgggacggcg 1658

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hg.*
- 3: gb_in.*
- 4: gb_ov.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_or.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	25	1.5	262181	2	AC125659	AC125659 Rattus no
2	25	1.5	277282	2	AC137423	AC137423 Rattus no
3	23	1.4	173222	9	AC009127	AC009127 Homo sapi
4	23	1.4	174933	9	AC009087	AC009087 Homo sapi
5	23	1.4	184391	2	AC115277	AC115277 Rattus no
6	22	1.3	592	4	AY285636	AY285636 Sus scrofa
7	22	1.3	2409	6	AX653883	AX653883 Sequence
8	22	1.3	109964	2	AP005677	AP005677 Oryza sat
9	22	1.3	150876	2	AC108761	AC108761 Oryza sat
10	22	1.3	151468	2	AC116118	AC116118 Mus muscu
11	22	1.3	176045	10	AC122370	AC122370 Mus muscu
12	22	1.3	191735	10	AL596446	AL596446 Mouse DNA
13	22	1.3	197694	2	AC136065	AC136065 Rattus no
14	22	1.3	197981	2	AC117089	AC117089 Rattus no
15	22	1.3	215867	2	AC099439	AC099439 Rattus no
16	22	1.3	223363	2	AC120700	AC120700 Rattus no
17	22	1.3	242309	2	AC098167	AC098167 Rattus no
18	22	1.3	245893	2	AC125267	AC125267 Mus muscu
19	22	1.3	256105	2	AC093979	AC093979 Rattus no
20	22	1.3	261414	2	AC122960	AC122960 Rattus no
21	22	1.3	277860	2	AC124647	AC124647 Mus muscu
22	22	1.3	290002	2	AC134106	AC134106 Rattus no
23	22	1.3	294415	2	AC110321	AC110321 Rattus no
24	21	1.3	426	8	AF218307	AF218307 Saccharom
25	21	1.3	802	8	AY089172	AY089172 Arabidops
26	21	1.3	3717	1	D49692	D49692 Spirulina p
27	21	1.3	5956	6	AX251973	AX251973 Sequence
28	21	1.3	5956	6	AX344365	AX344365 Sequence
29	21	1.3	5956	6	AX348758	AX348758 Sequence
30	21	1.3	11915	1	AE009922	AE009922 Pyrobacul
31	21	1.3	28573	3	AF024502	AF024502 Caenorhab
32	21	1.3	31219	2	AC019857	AC019857 Drosophil
33	21	1.3	39000	9	AP003024	AP003024 Homo sapi
34	21	1.3	69646	2	AC102089	AC102089 Mus muscu
35	21	1.3	73746	2	AC101286	AC101286 Mus muscu
36	21	1.3	93216	9	AC003016	AC003016 Human BAC
37	21	1.3	98535	9	AF188025	AF188025 Homo sapi
38	21	1.3	99150	9	AF202964	AF202964 Homo sapi
39	21	1.3	112070	9	AC068798	AC068798 Homo sapi
40	21	1.3	119654	2	AC134325	AC134325 Mus muscu
41	21	1.3	129779	3	AC006214	AC006214 Drosophil
42	21	1.3	134784	8	FL17123	FL17123 Arabidops
43	21	1.3	143441	2	AC013369	AC013369 Homo sapi
44	21	1.3	144290	9	AC034251	AC034251 Homo sapi
45	21	1.3	146811	9	AF121898	AF121898 Homo sapi

ALIGNMENTS

RESULT 1
AC125659
LOCUS
DEFINITION Rattus norvegicus clone CH230-11B1, WORKING DRAFT SEQUENCE.
AC125659
ACCESSION
VERSION HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 262181)
Muzny, D.Marie., Metzker, M.Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

2

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gregorogreis, E., Geet, K., Gill, R., Grady, A., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, J., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoques, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, D., Kowar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Loulseghe, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Mallory, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mahoney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munitasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackelme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, N., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, J., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, I., Thomas, R., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Unpublished
Direct Submission
2 (bases 1 to 277282)
Rat Genome Sequencing Consortium.

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Genome Center
Center: Baylor College of Medicine

Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: KZLE

Center clone name: CH230-unknown
----- Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 204050 bases at least Q40

Consensus quality: 209444 bases at least Q30
Consensus quality: 212955 bases at least Q20
Estimated insert size: 204859; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 12348: contig of 12348 bp in length
* 12349 12448: gap of unknown length
* 12449 251860: contig of 239412 bp in length
* 251861 251960: gap of unknown length
* 251961 255729: contig of 3769 bp in length
* 255730 255830: gap of unknown length
* 255831 269031: contig of 13102 bp in length
* 269032 27015: gap of unknown length
* 27016 270815: gap of unknown length
* 270816 274241: gap of unknown length
* 274242 277282: contig of 2942 bp in length.
* 274341

FEATURES
source

1. 277282
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-unknown"
42559. 44206
/note="wgs_contig"
97204. .99317
/note="wgs_contig"
115739. .118162
/note="wgs_contig"
255830. 257305
/note="wgs_contig"
268832. .268931
/note="wgs_contig"

ORIGIN

Query Match 1.5% Score 25; DB 2; Length 277282;
Best Local Similarity 100.0%; Pred.No. 0.17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1167 CTGCCCCAAGAAAGGTTTCTTAACA 1191
|||||
DB 84796 CTGCCCCAAGAAAGGTTTCTTAACA 84820
|||||

RESULT 3
AC009127/c

LOCUS AC009127 173222 bp DNA linear PRI 12-APR-2002
DEFINITION Homo sapiens chromosome 16 clone RP11-498D10, complete sequence.
ACCESSION AC009127
VERSION AC009127.9 GI:20143529
KEYWORDS HTG.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 173222)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 173222)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission

JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 3 (bases 1 to 173222)

AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission

JOURNAL Submitted (12-APR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT On Apr 12, 2002 this sequence version replaced gi:13786311.
Sequence Quality Assessment:
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

Sequence Quality Assessment:
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

FEATURES
source
1. Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-498D10"

ORIGIN
Query Match 1.4%; Score 23; DB 9; Length 173222;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1546 CAAAGCAGCAGCACAGCCCAACA 1568
|||||
Db 45290 CAAAGCAGCAGCACAGCCCAACA 45268

RESULT 4
AC009087/c
LOCUS AC009087 174933 bp DNA linear PRI 27-APR-2000
DEFINITION Homo sapiens chromosome 16 clone RP11-384M15, complete sequence.
ACCESSION AC009087
VERSION AC009087.4 GI:7656694
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 174933)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 174933)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 174933)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Apr 27, 2000 this sequence version replaced gi:7025737.
Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.7% of Sequence;
Estimated Total Number of Errors is 1.
STS Content:
SHGC-32168 G217137
WI-15979 G21275
WI-10173 G11750
WI-17126 G21504
WI-9359 G06054.

FEATURES
1. Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-384M15"

ORIGIN
Query Match 1.4%; Score 23; DB 9; Length 174933;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1546 CAAAGCAGCAGCACAGCCCAACA 1568
|||||
Db 150280 CAAAGCAGCAGCACAGCCCAACA 150258

RESULT 5
AC115277/c
LOCUS AC115277 184391 bp DNA linear HTG 15-NOV-2002
DEFINITION Rattus norvegicus clone CH230-372C1, *** SEQUENCING IN PROGRESS ***
ACCESSION AC115277
VERSION AC115277.5 GI:25009102
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 184391)
AUTHORS Muzny, D., Metzker, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooke, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, A., Chu, J., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geier, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hollins, B., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hernandez, R., Hines, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Loresuhewa, L., Louised, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelen, O., Okwunodu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfamkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Unpublished
2 (bases 1 to 184391)
Worley, K.C.
Direct Submission
Submitted (16-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 184391)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23194609.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GTHP
Center clone name: CH230-372C1
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 164483 bases at least Q40
Consensus quality: 167335 bases at least Q30
Consensus quality: 169474 bases at least Q20
Estimated insert size: 170619; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

*
FEATURES
source

1 184391: contig of 184391 bp in length.
Location/Qualifiers
1. 184391
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-372C1"
1. 1601
/note="wgs end extension
clone end:Sp6"
1705_3257
/note="wgs end extension
clone end:Sp6"
4527_107311
/note="clone boundary
clone end:Sp6
site:
end sequence:BZ216963"
170821_173036
/note="wgs contig"
complement(182070_182951)
/note="clone boundary
clone end:T7
site:
end sequence:BZ216960"

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

ORIGIN

Query Match 1.4%; Score 23; DB 2; Length 184391;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 842 TTTGTGCTTTTCATGTATGTGCT 864
|||||
Db 77557 TTTGTGCTTTTCATGTATGTGCT 77535
|||||

RESULT 6

AY285636/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

repeat_region

ORIGIN

Query Match

Best Local Similarity

Matches

Qy

Db

RESULT 7

AY285636

Sus scrofa

Sus scrofa

Sus scrofa

Sus scrofa

Sus scrofa

Sus scrofa

Sus scrofa

Sus scrofa

Sus scrofa

Sus scrofa

Sus scrofa

Sus scrofa

Sus scrofa

```

AX653883/c
LOCUS AX653883 2409 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 3753 from Patent WO03000898.
ACCESSION AX653883
VERSION AX653883.1 GI:29156697
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 Chang H.S., Chen W., Cooper B., Glazebrook J., Goff S.A., Hou Y.M.,
Katagiri F., Qian S., Tao Y., Whitham S., Xie Z., Zhu T. and Zou G.
Plant genes involved in defense against pathogens
Patent: WO 03000898-A 3753 03-JAN-2003;
Syngenta Participations AG (CH)
LOCATION/Qualifiers
1. .2409
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"
ORIGIN
Query Match 1.3%; Score 22; DB 6; Length 2409;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 657 CGACGACGTCGACGCTGATGCT 678
|||||
DB 492 CGACGACGTCGACGCTGATGCT 471

RESULT 8
AP005677/c
LOCUS AP005677 109964 bp DNA linear HTG 06-SEP-2002
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 9 clone
OJ1011 C06, *** SEQUENCING IN PROGRESS ***
ACCESSION AP005677
VERSION AP005677.1 GI:22775435
HTG: HTGS PHASE2.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 Sasaki, T., Matsumoto, T. and Katayose, Y.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, BAC
clone: OJ1011 C06
Published Only in Database (2002)
JOURNAL
REFERENCE
2 (bases 1 to 109964)
AUTHORS Sasaki, T., Matsumoto, T. and Katayose, Y.
TITLE Direct Submission
JOURNAL Submitted (05-SEP-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
The nucleotide sequence of this BAC clone was generated by
combining Monsanto and RGP-Japan sequencing data.
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
LOCATION/Qualifiers
1. .109964
FEATURES
source

```

```

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="9"
/clone="OJ1011_C06"
ORIGIN
Query Match 1.3%; Score 22; DB 2; Length 109964;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 657 CGACGACGTCGACGCTGATGCT 678
|||||
DB 77096 CGACGACGTCGACGCTGATGCT 77075

RESULT 9
AC108761/c
LOCUS AC108761 150876 bp DNA linear HTG 31-JAN-2002
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 9 clone
OSUNBa0087J09, *** SEQUENCING IN PROGRESS ***; 7 unordered pieces.
ACCESSION AC108761
VERSION AC108761.1 GI:18449966
HTG: HTGS PHASE1.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 (bases 1 to 150876)
AUTHORS Lee, J.-S., Hahn, J.-H., Yoon, U.-H., Yun, D.-W., Lee, M.-C., Eun, M.Y.
and Kim, H.-I.
TITLE Oryza sativa BAC OSUNBa0087J09 genomic sequence
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 150876)
AUTHORS Hahn, J.-H. and Kim, H.-I.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2002) Rice Genome Sequencing Project, National
Institute of Agricultural Science and Technology (NIAS), RDA, 249
Seodun-dong, Suwon 441-707, Korea (E-mail: jhahn@rda.go.kr,
Tel: 82-31-290-0309, Fax: 82-31-290-0308)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 45839: contig of 45839 bp in length
* 45840 45939: gap of unknown length
* 45940 76833: contig of 30894 bp in length
* 76834 76934: gap of unknown length
* 76934 109205: contig of 32272 bp in length
* 109205 109305: gap of unknown length
* 109305 132339: contig of 22934 bp in length
* 132339 132340: gap of unknown length
* 132340 134166: contig of 1827 bp in length
* 134167 134266: gap of unknown length
* 134267 140337: contig of 5971 bp in length
* 140338 140337: gap of unknown length
* 140338 150876: contig of 10539 bp in length.
LOCATION/Qualifiers
1. .150876
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/variety="Nipponbare"
/db_xref="taxon:39947"
/chromosome="9"
/clone="OSUNBa0087J09"
ORIGIN
FEATURES
source

```

Query Match 1.3%; Score 22; DB 2; Length 150876;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 657 CGACGACGTCGACGCTGATGCT 678
 DB 118775 CGACGACGTCGACGCTGATGCT 118754

RESULT 10
 AC116118
 LOCUS
 DEFINITION Mus musculus chromosome 16 clone RP23-73P19 map 16, *** SEQUENCING
 IN PROGRESS ***, 8 unordered pieces.
 AC116118
 AC116118.5 GI:39841503
 HTGS; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
 Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 151468)
 Birren, B., Nusbaum, C. and Lander, E.
 Mus musculus chromosome 16, clone RP23-73P19
 Unpublished
 2 (bases 1 to 151468)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collamore, A.,
 Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S.,
 Faro, S., Ferreira, P., FitzHugh, M., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, M., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kellis, C., Larocque, K., Lamazares, R.,
 Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 MacLean, C., MacDonald, P., Major, J., Marcus, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L.,
 Mihova, T., Mieng, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission

TITLE
 JOURNAL
 REFERENCE
 AUTHORS

Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 151468)
 Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
 Collamore, A., Cooke, P., Corum, B., DeArelano, K., Faro, S.,
 Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S.,
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hagos, B., Hagopian, D., Hagos, B.,
 Hall, J., Horton, L., Hulme, M., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C.,
 MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrum, J., Menus, L., Mihova, T., Mieng, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission

TITLE

JOURNAL
 COMMENT
 Submitted (15-DEC-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Dec 15, 2003 this sequence version replaced gi:28460918.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: h23300
 Center clone name: 73_P19

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 8 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 13387: contig of 13387 bp in length
 * 13388 13487: gap of 100 bp
 * 13488 18518: contig of 5031 bp in length
 * 18519 18618: gap of 100 bp
 * 18619 35043: contig of 16425 bp in length
 * 35044 35143: gap of 100 bp
 * 35144 54747: contig of 19604 bp in length
 * 54748 54847: gap of 100 bp
 * 54848 64657: contig of 9810 bp in length
 * 64658 64758: gap of 100 bp
 * 64758 101040: contig of 36283 bp in length
 * 101041 101140: gap of 100 bp
 * 101141 130194: contig of 23054 bp in length
 * 130195 130294: gap of 100 bp
 * 130295 151468: contig of 21174 bp in length.

FEATURES
 source
 1. 151468
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="16"
 /map="16"
 /clone_lib="RP23-73P19"
 /clone="RP23-73P19"

ORIGIN
 Query Match 1.3%; Score 22; DB 2; Length 151468;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 TTTTGTGTTTCTGTTTCTGCTG 711
 DB 51549 TTTTGTGTTTCTGTTTCTGCTG 51570

RESULT 11
 AC122370
 LOCUS
 DEFINITION Mus musculus BAC clone RP23-45988 from chromosome 8, complete
 sequence.
 AC122370
 AC122370.4 GI:28376820
 HTGS
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 176045)
 Berghoff, A., Haakenson, W., Bielicki, L. and Meyer, R.
 The sequence of Mus musculus BAC clone RP23-45988


```

repeat_region 30802..31095
/rpt_family="L1"
repeat_region 31104..31287
/rpt_family="L1"
repeat_region 31663..33668
/rpt_family="L1"
repeat_region 33667..33893
/rpt_family="L1"
repeat_region 33893..34068
/rpt_family="L1"
repeat_region 34072..34262
/rpt_family="B2"
repeat_region 34282..34355
/rpt_family="L1"
repeat_region 35311..35384
/rpt_family="MaLR"
repeat_region 35526..35938
/rpt_family="MaLR"
repeat_region 35976..36243
/rpt_family="L1"
repeat_region 36413..37025
/rpt_family="L1"
repeat_region 37081..37528
/rpt_family="L1"
repeat_region 37668..37812
/rpt_family="L1"
repeat_region 37893..38061
/rpt_family="Alu"
repeat_region 38203..38420
/rpt_family="B4"
repeat_region 39026..39165
/rpt_family="L1"
repeat_region 40708..40898
/rpt_family="MaLR"
repeat_region 40947..41307
/rpt_family="L1"
repeat_region 41599..41691
/rpt_family="Alu"
repeat_region 43021..43288
/rpt_family="L1"
repeat_region 43259..43417

```

Query Match 1.3%; Score 22; DB 10; Length 176045;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 TATTTTCTTAAAGAAAAA 248
 Db 88207 TATTTTCTTAAAGAAAAA 88228

RESULT 12
 AC1596446
 LOCUS
 DEFINITION Mouse DNA sequence from clone RP23-386E10 on chromosome 11,
 complete sequence.
 ACCESSION AL596446
 VERSION AL596446.11 GI:20068511
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Tracey, A.
 Direct Submission
 Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk
 On Apr 7, 2002 this sequence version replaced gi:17043843.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em: EMBL; SW: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP
 database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-386E10 is
 from the RPCI-23 Mouse PAC Library
 constructed by the group of Pieter de Jong.

For further details see http://www.chori.org/bacpac/home.htm

VECTOR: pBAC3.6.

FEATURES

Location/Qualifiers
 1..191735
 /organism="Mus musculus"
 /mol_type="Genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="11"
 /clone="RP23-386E10"
 /clone_lib="RPCI-23"

ORIGIN

Query Match 1.3%; Score 22; DB 10; Length 191735;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 TTTTCTTTTCTCTCTCTG 711
 Db 157107 TTTTCTTTTCTCTCTG 157128

RESULT 13
 AC136065/c

LOCUS
 AC136065 197694 bp DNA linear HTG 29-OCT-2002
 Rattus norvegicus clone CH230-85M19, *** SEQUENCING IN PROGRESS
 ***, 85 unordered pieces.

AC136065
 HTG: HTGS PHASE1
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

1 (bases 1 to 197694)

REFERENCE

AUTHORS
 Muzny, D., Maric, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Bialwalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, S.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.W., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gubaregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevata, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hoques, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

10087	contig of 1097 bp in length	8987	contig of 1097 bp in length
10088	gap of unknown length	8988	gap of unknown length
10089	contig of 1348 bp in length	8989	contig of 1348 bp in length
10090	gap of unknown length	8990	gap of unknown length
10091	contig of 1057 bp in length	8991	contig of 1057 bp in length
10092	gap of unknown length	8992	gap of unknown length
10093	contig of 1304 bp in length	8993	contig of 1304 bp in length
10094	gap of unknown length	8994	gap of unknown length
10095	contig of 1562 bp in length	8995	contig of 1562 bp in length
10096	gap of unknown length	8996	gap of unknown length
10097	contig of 1203 bp in length	8997	contig of 1203 bp in length
10098	gap of unknown length	8998	gap of unknown length
10099	contig of 1212 bp in length	8999	contig of 1212 bp in length
10100	gap of unknown length	9000	gap of unknown length
10101	contig of 1612 bp in length	9001	contig of 1612 bp in length
10102	gap of unknown length	9002	gap of unknown length
10103	contig of 1479 bp in length	9003	contig of 1479 bp in length
10104	gap of unknown length	9004	gap of unknown length
10105	contig of 1263 bp in length	9005	contig of 1263 bp in length
10106	gap of unknown length	9006	gap of unknown length
10107	contig of 1656 bp in length	9007	contig of 1656 bp in length
10108	gap of unknown length	9008	gap of unknown length
10109	contig of 1657 bp in length	9009	contig of 1657 bp in length
10110	gap of unknown length	9010	gap of unknown length
10111	contig of 1709 bp in length	9011	contig of 1709 bp in length
10112	gap of unknown length	9012	gap of unknown length
10113	contig of 1045 bp in length	9013	contig of 1045 bp in length
10114	gap of unknown length	9014	gap of unknown length
10115	contig of 1919 bp in length	9015	contig of 1919 bp in length
10116	gap of unknown length	9016	gap of unknown length
10117	contig of 2586 bp in length	9017	contig of 2586 bp in length
10118	gap of unknown length	9018	gap of unknown length
10119	contig of 2553 bp in length	9019	contig of 2553 bp in length
10120	gap of unknown length	9020	gap of unknown length
10121	contig of 1521 bp in length	9021	contig of 1521 bp in length
10122	gap of unknown length	9022	gap of unknown length
10123	contig of 1961 bp in length	9023	contig of 1961 bp in length
10124	gap of unknown length	9024	gap of unknown length
10125	contig of 1408 bp in length	9025	contig of 1408 bp in length
10126	gap of unknown length	9026	gap of unknown length
10127	contig of 1593 bp in length	9027	contig of 1593 bp in length
10128	gap of unknown length	9028	gap of unknown length
10129	contig of 2033 bp in length	9029	contig of 2033 bp in length
10130	gap of unknown length	9030	gap of unknown length
10131	contig of 1813 bp in length	9031	contig of 1813 bp in length
10132	gap of unknown length	9032	gap of unknown length
10133	contig of 2397 bp in length	9033	contig of 2397 bp in length
10134	gap of unknown length	9034	gap of unknown length
10135	contig of 2792 bp in length	9035	contig of 2792 bp in length
10136	gap of unknown length	9036	gap of unknown length
10137	contig of 3362 bp in length	9037	contig of 3362 bp in length
10138	gap of unknown length	9038	gap of unknown length
10139	contig of 1789 bp in length	9039	contig of 1789 bp in length
10140	gap of unknown length	9040	gap of unknown length
10141	contig of 1897 bp in length	9041	contig of 1897 bp in length
10142	gap of unknown length	9042	gap of unknown length
10143	contig of 2414 bp in length	9043	contig of 2414 bp in length
10144	gap of unknown length	9044	gap of unknown length
10145	contig of 2220 bp in length	9045	contig of 2220 bp in length
10146	gap of unknown length	9046	gap of unknown length
10147	contig of 2268 bp in length	9047	contig of 2268 bp in length
10148	gap of unknown length	9048	gap of unknown length
10149	contig of 2226 bp in length	9049	contig of 2226 bp in length
10150	gap of unknown length	9050	gap of unknown length
10151	contig of 3151 bp in length	9051	contig of 3151 bp in length
10152	gap of unknown length	9052	gap of unknown length
10153	contig of 2274 bp in length	9053	contig of 2274 bp in length
10154	gap of unknown length	9054	gap of unknown length
10155	contig of 2727 bp in length	9055	contig of 2727 bp in length
10156	gap of unknown length	9056	gap of unknown length
10157	contig of 2719 bp in length	9057	contig of 2719 bp in length
10158	gap of unknown length	9058	gap of unknown length
10159	contig of 3629 bp in length	9059	contig of 3629 bp in length
10160	gap of unknown length	9060	gap of unknown length

Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorussehwa, L., Lou, H., Lozado, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartine, M., Mahmoud, M., Mallory, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawney, S., Mcleod, M., McNeill, T., Meenen, E., Milosavljevic, A.,
 Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K.,
 Morris, S., Munkidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D.,
 Newton, N., Nguyen, N., Norris, S., Nwaokemeleh, O., Okwuonu, G.,
 Olarunpisaqon, A., Pal, S., Parks, K., Pasternak, S., Paul, H.,
 Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A.,
 Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E.,
 Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,
 Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A.,
 Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S.,
 Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A.,
 Sisson, I., Sitter, C.D., Smajd, D., Sneed, A., Sodergren, E.,
 Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A.,
 Svatek, A., Tabor, P., Taylor, C., Taylor, R., Thomas, N., Thomas, S.,
 Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D.,
 Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J.,
 Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R.,
 Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,
 Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X.,
 Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R.,
 Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.
 Direct Submission
 2 (bases 1 to 197694)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (29-OCT-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: KDBI
 Center clone name: CH230-85M19
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 146097 bases at least Q40
 Consensus quality: 151106 bases at least Q30
 Consensus quality: 155140 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 65 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1
 * 1241: contig of 1241 bp in length
 * 1242: gap of unknown length
 * 1342: contig of 1510 bp in length
 * 2852: gap of unknown length
 * 2952: contig of 1221 bp in length
 * 4173: gap of unknown length
 * 4272: contig of 1402 bp in length
 * 5674: gap of unknown length
 * 5774: gap of unknown length
 * 5775: contig of 1518 bp in length
 * 7293: gap of unknown length
 * 7393: contig of 1398 bp in length
 * 8791: gap of unknown length

```

* 84947 85046: gap of unknown length
* 85047 88159: contig of 3113 bp in length
* 88160 88259: gap of unknown length
* 88260 92104: contig of 3845 bp in length
* 92105 92204: gap of unknown length
* 92205 96596: contig of 4392 bp in length
* 96597 96696: gap of unknown length
* 96697 100965: contig of 4269 bp in length
* 100966 101065: gap of unknown length
* 101066 104609: contig of 3544 bp in length
* 104610 109250: contig of 4541 bp in length
* 109251 109350: gap of unknown length
* 109351 113836: contig of 4486 bp in length
* 113837 118366: gap of unknown length
* 118367 118684: contig of 4648 bp in length
* 118685 122698: contig of 4014 bp in length
* 122699 127266: contig of 4468 bp in length
* 127267 127366: gap of unknown length

Query Match 1.3%; Score 22; DB 2; Length 197694;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 225 TCTATTTTCTTAAAGAAAAA 246
Db 40901 TCTATTTTCTTAAAGAAAAA 40880

RESULT 14
AC117089
LOCUS
DEFINITION
Rattus norvegicus clone CH230-11969, WORKING DRAFT SEQUENCE, 2
unordered pieces.
AC117089
AC117089.4 GI:25008560
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 197981)
Mazny D.Marie., Metzker M.Lee., Abramson S., Adams C., Alder J.,
Allen C., Allen H., Albrooks S., Amin A., Anguiano D.,
Anyalebechi V., Aoyagi A., Ayodeji M., Baca E., Baden H.,
Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,
Biswal N., Blair J., Blankenburg K., Blyth P., Brown M.,
Bryant N., Buhay C., Burch P., Burrell K., Calderon E.,
Cardenas V., Carter K., Cavazos I., Ceasar H., Center A.,
Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J.,
Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L.,
Davila M.L., Davis C., Davy-Carroll L., De Anda C., Dederich D.,
Delgado O., Denison S., Detamo C., Ding Y., Dinh H., Divya K.,
Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K.,
Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G.,
Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,
Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M.,
Gebregeorgis E., Geer K., Gill R., Grady M., Guerra W., Guevara W.,
Gunnaratne P., Haaland W., Hamil C., Hamilton C., Hamilton K.,
Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez J.,
Hernandez R., Hines S., Hladun S.L., Hodgson A., Hogues M.,
Hollins B., Howells S., Hulyk S., Hume J., Idlebird D., Jackson A.,
Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolivet A.,
Karpathy S., Kelly S., Khan Z., King L., Kovar C.,
Kowis C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J.,
Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J.,
Lorenshewa L., Loulseged H., Lozado R.J., Lu X., Ma J.,
Maheshwari M., Mahindran M., Mahmoud M., Malloy K., Mangum A.,
Mangum B., Mapua P., Martin K., Martin R., Martinez E.,
Mawhinney S., McLeod M.P., McNeill T.Z., Meenen E.,

```

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munitasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelameh, O., Okwuonu, G., Olarpunsgoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, D., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
TITLE
REFERENCE
AUTHORS
JOURNAL
COMMENT

Direct Submission
Unpublished
2 (bases 1 to 197981)
Worley, K.C.
Direct Submission
Submitted (08-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 197981)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23194925.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Genome Center
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GFPD
Center clone name: CH230-11969
Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 189549 bases at least Q40
Consensus quality: 190935 bases at least Q30
Consensus quality: 192081 bases at least Q20
Estimated insert size: 196342; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will

* be preserved.
 * 1 196422: contig of 196422 bp in length
 * 196423 196522: gap of unknown length
 * 196523 197981: contig of 1459 bp in length.
 FEATURES
 Source
 misc_feature
 194929. 195819
 /note="clone_boundary
 clone_end:T7
 site:
 end_sequence:BH315129"

ORIGIN

Query Match 1.3%; Score 22; DB 2; Length 197981;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 690 TTTTGTGTTTGTCTCTCTGTG 711
 |||||
 Db 53534 TTTTGTGTTTGTCTCTCTGTG 53555

RESULT 15

AC099439/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-145P16, WORKING DRAFT SEQUENCE.
 AC099439
 AC099439.5 GI:30521297
 HTG: HTGS_PHASE2; HTGS DRAFT; HTGS_FULLTOP.
 VERSION
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS

1 (bases 1 to 215867)
 Muzny, D., Marie, E., Metzker, M., Lee, S., Adams, C., Alder, J., Allen, C., Allen, H., Altschuler, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diyya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Georegeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idiebird, D., Jackson, A., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuewa, L., Louised, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mahoney, S., McLeod, M.P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokeleleh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L.,

Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmami, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhao, S., Zhou, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 215867)
 Worley, K.C.
 Direct Submission
 Submitted (15-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 215867)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On May 10, 2003 this sequence version replaced gi:22855268.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GIHG
 Center clone name: CH230-145P16
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 210812 bases at least Q40
 Consensus quality: 212267 bases at least Q30
 Consensus quality: 213291 bases at least Q20
 Estimated insert size: 221076; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 215867: contig of 215867 bp in length.
 * Location/Qualifiers
 1. 215867
 /organism="Rattus norvegicus"

FEATURES
 source

```

/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-145P16"
misc_feature 1..1660
/note="wgs_contig"

```

ORIGIN

```

Query Match      1.3%; Score 22; DB 2; Length 215867;
Best Local Similarity 100.0%; Pred.No. 7.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      690 TTTGTTTGTGTTTCTCTGTG 711
          |||||||
Db      46567 TTTGTTTGTGTTTCTCTGTG 46546

```

Search completed: September 2, 2004, 20:00:48
Job time : 19151 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 10:43:49 ; Search time 2015 Seconds
(without alignments)
3495.540 Million cell updates/sec

Title: US-10-732-721-1

Perfect score: 1658

Sequence: 1 gaattacggctcacatac.....ctcgttagttggacggcg 1658

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq_29Jan04.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002as.*
- 7: geneseqn2003as.*
- 8: geneseqn2003bs.*
- 9: geneseqn2003cs.*
- 10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	22	1.3	2409	7	ADA70430 Rice gene
C 2	21	1.3	801	3	AAC50800 Arabidops
C 3	21	1.3	802	3	AAC33054 Arabidops
C 4	21	1.3	5956	6	ABK31369 Signal tr
C 5	21	1.3	5956	6	ABL70326 Chemical
C 6	21	1.3	5956	6	AAS61273 Human gen
C 7	21	1.3	16739	4	ABL07848 Drosophil
C 8	21	1.3	21828	4	ABL07850 Drosophil
C 9	20	1.2	291	6	ABN17338 Human ORF
C 10	20	1.2	785	4	AAL15883 Human bre
C 11	20	1.2	830	8	ADA49079 Wheat gen
C 12	20	1.2	1044	7	ACC61661 Gene sequ
C 13	20	1.2	1963	3	AAZ52423 HTRM clon
C 14	20	1.2	1973	4	AAL60732 Human pol
C 15	20	1.2	1979	4	AAL58946 Human pol
C 16	20	1.2	1979	8	ADB48928 Novel hum
C 17	20	1.2	3255	7	ADA53534 Human cod
C 18	20	1.2	4633	4	AAL93684 Human TNF
C 19	20	1.2	5823	6	ABK28382 DNA trans
C 20	20	1.2	12988	4	AAK80680 Human imm
C 21	20	1.2	12988	7	ABZ74054 Secreted
C 22	20	1.2	12988	9	ADC20775 Human sec
C 23	20	1.2	16163	4	AAL62620 Human bre

24	20	1.2	16163	4	AAK83901	Aak83901 Human imm
25	20	1.2	16163	4	AAK87161	Aak87161 Human imm
c 26	20	1.2	16163	4	AAL03817	Aal03817 Human rep
27	20	1.2	100301	6	ABQ88176	Abq88176 Human ost
c 28	19	1.1	558	3	AAF08066	Aaf08066 Fusarium
c 29	19	1.1	604	6	ABN87803	Abn87803 Human ova
c 30	19	1.1	648	7	ABZ54983	Abz54983 Aspergill
c 31	19	1.1	734	6	ABK85714	Abk85714 DNA encod
c 32	19	1.1	771	4	AAK59627	Aak59627 Human imm
c 33	19	1.1	911	6	ABK99945	Abk99945 DNA encod
c 34	19	1.1	1087	4	AAK75762	Aak75762 Human imm
c 35	19	1.1	1447	2	AAV08179	Aav08179 Bovine in
c 36	19	1.1	1447	2	AAV68162	Aav68162 Bovine in
c 37	19	1.1	1449	1	AAN30041	Aan30041 Sequence
c 38	19	1.1	1449	6	ABS53995	Abs53995 cDNA enco
c 39	19	1.1	1469	1	AAN30039	Aan30039 Sequence
c 40	19	1.1	1469	2	AAV08177	Aav08177 Bovine in
c 41	19	1.1	1469	2	AAV68160	Aav68160 Bovine in
c 42	19	1.1	1469	6	ABS53993	Abs53993 cDNA enco
c 43	19	1.1	1484	9	ADD30329	Add30329 Plant yie
c 44	19	1.1	1551	6	ABZ31860	Abz31860 Candida a
c 45	19	1.1	1557	6	ABQ67825	Abq67825 Listeria

ALIGNMENTS

RESULT 1

ADA70430/c	
ID	ADA70430 standard; DNA; 2409 BP.
AC	ADA70430;
XX	
DT	20-NOV-2003 (first entry)
XX	
DE	Rice gene, SEQ ID 3753.
XX	
KW	Plant; bacterial infection; fungal infection; viral infection; rice;
KW	gene; ds.
XX	
OS	Oryza sativa.
XX	
PN	WO2003000898-A1.
XX	
FD	03-JAN-2003.
XX	
PF	22-JUN-2001; 2001WO-IB001105.
XX	
PR	22-JUN-2001; 2001WO-IB001105.
XX	
FA	(SYGN) SYNGENTA PARTICIPATIONS AG.
XX	
PI	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI	Katagiri P, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX	
DR	WPI; 2003-175296/17.
XX	
PT	Identifying at least one gene involved in plant resistance or response to
PT	pathogenic infection for conferring resistance or tolerance to a plant to
PT	bacterial, fungal or viral infection by determining or detecting plant
PT	gene expression.
XX	
FS	Claim 6; SEQ ID NO 3753; 899pp; English.
XX	

The present invention relates to a method (M1) for identifying genes involved in plant resistance or response to pathogenic infection. M1 comprises identifying a gene whose expression is significantly altered in the incompatible interaction of plant gene expression relative to expression of the gene in an uninfected plant, in a mutant plant that does not express a gene associated with response to pathogenic infection, or in a corresponding incompatible or compatible interaction. (M1) is useful for conferring resistance to resistance or tolerance to a plant to bacterial, fungal or viral infection. The present sequence was used to

```
CC illustrate the invention.
XX
SQ Sequence 2409 BP; 399 A; 846 C; 799 G; 361 T; 0 U; 4 Other;
    Query Match      1.3%; Score 22; DB 7; Length 2409;
    Best Local Similarity 100.0%; Pred. No. 3.5;
    Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 657 CGACGACGTCGACGCTGATGCT 678
Db 492 CGACGACGTCGACGCTGATGCT 471

RESULT 2
AAC50800/c
ID AAC50800 standard; DNA; 801 BP.
AC
XX AAC50800;
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 66182.
XX
DE Hybridisation assay; Genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR
07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140821P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
```


PR 06-AUG-1999;	99US-0147416P.	Qy	691	TTTGTCTTTTCTCTCTGTG 711
PR 09-AUG-1999;	99US-0147433P.	Db	24	TTTTTTTTTTTTTTTTTTTT
PR 09-AUG-1999;	99US-0147935P.			TTTTTTTTTTTTTTTTTTTT
PR 10-AUG-1999;	99US-0148171P.			TTTTTTTTTTTTTTTTTTTT
PR 11-AUG-1999;	99US-0148319P.			TTTTTTTTTTTTTTTTTTTT
PR 12-AUG-1999;	99US-0148341P.			TTTTTTTTTTTTTTTTTTTT
PR 13-AUG-1999;	99US-0148565P.			TTTTTTTTTTTTTTTTTTTT
PR 13-AUG-1999;	99US-0148684P.			TTTTTTTTTTTTTTTTTTTT
PR 16-AUG-1999;	99US-0149368P.			TTTTTTTTTTTTTTTTTTTT
PR 17-AUG-1999;	99US-0149175P.			TTTTTTTTTTTTTTTTTTTT
PR 18-AUG-1999;	99US-0149426P.			TTTTTTTTTTTTTTTTTTTT
PR 20-AUG-1999;	99US-0149723P.			TTTTTTTTTTTTTTTTTTTT
PR 20-AUG-1999;	99US-0149722P.			TTTTTTTTTTTTTTTTTTTT
PR 20-AUG-1999;	99US-0149929P.			TTTTTTTTTTTTTTTTTTTT
PR 23-AUG-1999;	99US-0149902P.			TTTTTTTTTTTTTTTTTTTT
PR 23-AUG-1999;	99US-0149930P.			TTTTTTTTTTTTTTTTTTTT
PR 25-AUG-1999;	99US-0150566P.			TTTTTTTTTTTTTTTTTTTT
PR 26-AUG-1999;	99US-0150884P.			TTTTTTTTTTTTTTTTTTTT
PR 27-AUG-1999;	99US-0151065P.			TTTTTTTTTTTTTTTTTTTT
PR 27-AUG-1999;	99US-0151066P.			TTTTTTTTTTTTTTTTTTTT
PR 27-AUG-1999;	99US-0151080P.			TTTTTTTTTTTTTTTTTTTT
PR 30-AUG-1999;	99US-0151303P.			TTTTTTTTTTTTTTTTTTTT
PR 31-AUG-1999;	99US-0151438P.			TTTTTTTTTTTTTTTTTTTT
PR 01-SEP-1999;	99US-0151930P.			TTTTTTTTTTTTTTTTTTTT
PR 07-SEP-1999;	99US-0152363P.			TTTTTTTTTTTTTTTTTTTT
PR 10-SEP-1999;	99US-0153070P.			TTTTTTTTTTTTTTTTTTTT
PR 13-SEP-1999;	99US-0153758P.			TTTTTTTTTTTTTTTTTTTT
PR 15-SEP-1999;	99US-0154018P.			TTTTTTTTTTTTTTTTTTTT
PR 16-SEP-1999;	99US-0154039P.			TTTTTTTTTTTTTTTTTTTT
PR 20-SEP-1999;	99US-0154779P.			TTTTTTTTTTTTTTTTTTTT
PR 22-SEP-1999;	99US-0155139P.			TTTTTTTTTTTTTTTTTTTT
PR 23-SEP-1999;	99US-0155486P.			TTTTTTTTTTTTTTTTTTTT
PR 24-SEP-1999;	99US-0155659P.			TTTTTTTTTTTTTTTTTTTT
PR 28-SEP-1999;	99US-0156458P.			TTTTTTTTTTTTTTTTTTTT
PR 29-SEP-1999;	99US-0156596P.			TTTTTTTTTTTTTTTTTTTT
PR 04-OCT-1999;	99US-0157117P.			TTTTTTTTTTTTTTTTTTTT
PR 05-OCT-1999;	99US-0157753P.			TTTTTTTTTTTTTTTTTTTT
PR 06-OCT-1999;	99US-0157865P.			TTTTTTTTTTTTTTTTTTTT
PR 07-OCT-1999;	99US-0158029P.			TTTTTTTTTTTTTTTTTTTT
PR 08-OCT-1999;	99US-0158232P.			TTTTTTTTTTTTTTTTTTTT
PR 12-OCT-1999;	99US-0158369P.			TTTTTTTTTTTTTTTTTTTT
PR 13-OCT-1999;	99US-0159293P.			TTTTTTTTTTTTTTTTTTTT
PR 13-OCT-1999;	99US-0159294P.			TTTTTTTTTTTTTTTTTTTT
PR 14-OCT-1999;	99US-0159329P.			TTTTTTTTTTTTTTTTTTTT
PR 14-OCT-1999;	99US-0159330P.			TTTTTTTTTTTTTTTTTTTT
PR 14-OCT-1999;	99US-0159331P.			TTTTTTTTTTTTTTTTTTTT
PR 14-OCT-1999;	99US-0159637P.			TTTTTTTTTTTTTTTTTTTT
PR 14-OCT-1999;	99US-0159638P.			TTTTTTTTTTTTTTTTTTTT
PR 18-OCT-1999;	99US-0159584P.			TTTTTTTTTTTTTTTTTTTT
PR 21-OCT-1999;	99US-0160741P.			TTTTTTTTTTTTTTTTTTTT
PR 21-OCT-1999;	99US-0160767P.			TTTTTTTTTTTTTTTTTTTT
PR 21-OCT-1999;	99US-0160768P.			TTTTTTTTTTTTTTTTTTTT
PR 21-OCT-1999;	99US-0160770P.			TTTTTTTTTTTTTTTTTTTT
PR 21-OCT-1999;	99US-0160814P.			TTTTTTTTTTTTTTTTTTTT
PR 21-OCT-1999;	99US-0160815P.			TTTTTTTTTTTTTTTTTTTT
PR 22-OCT-1999;	99US-0160900P.			TTTTTTTTTTTTTTTTTTTT
PR 22-OCT-1999;	99US-0160981P.			TTTTTTTTTTTTTTTTTTTT
PR 22-OCT-1999;	99US-0160989P.			TTTTTTTTTTTTTTTTTTTT
PR 25-OCT-1999;	99US-0161404P.			TTTTTTTTTTTTTTTTTTTT
PR 25-OCT-1999;	99US-0161405P.			TTTTTTTTTTTTTTTTTTTT
PR 26-OCT-1999;	99US-0161406P.			TTTTTTTTTTTTTTTTTTTT
PR 26-OCT-1999;	99US-0161359P.			TTTTTTTTTTTTTTTTTTTT
PR 26-OCT-1999;	99US-0161360P.			TTTTTTTTTTTTTTTTTTTT
PR 26-OCT-1999;	99US-0161361P.			TTTTTTTTTTTTTTTTTTTT
PR 28-OCT-1999;	99US-0161920P.			TTTTTTTTTTTTTTTTTTTT
PR 28-OCT-1999;	99US-0161922P.			TTTTTTTTTTTTTTTTTTTT
PR 28-OCT-1999;	99US-0161933P.			TTTTTTTTTTTTTTTTTTTT
PR 29-OCT-1999;	99US-0162142P.			TTTTTTTTTTTTTTTTTTTT

Query Match 1.3%; Score 21; DB 3; Length 801;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139817P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140353P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 28-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140991P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-0142390P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0145386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.

PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149823P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155653P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157573P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158023P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 1.3%; Score 21; DB 3; Length 802;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 691 TTGTTTTTGGTTTCTCTGTG 711
 Db 24 TTGTTTTTGGTTTCTCTGTG 4

RESULT 4

```
ABK31369
ID ABK31369 standard; DNA; 5956 BP.
XX
AC ABK31369;
XX
DT 23-APR-2002 (first entry)
XX
DE Signal transduction associated gene modified complementary DNA #106.
XX
KW Human; signal transduction associated gene; cytosine methylation state;
KW CpG island; signal transduction associated disease; solid tumour; cancer;
KW antitumour; cytostatic; mutant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX WO200200926-A2.
XX
XX 03-JAN-2002.
XX
XX 29-JUN-2001; 2001WO-EP007472.
XX
XX 30-JUN-2000; 2000DE-01032529.
XX
XX 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2002-147896/19.
XX
XX Oligonucleotide for diagnosis and therapy of diseases associated with
XX signal transduction e.g. cancer, comprises chemically modified genomic
XX sequences of genes associated with signal transduction.
XX
XX Claim 1; SEQ ID NO 212; 24pp; English.
XX
XX The present invention relates to chemically modified DNA sequences of
XX signal transduction associated genes. The DNA sequences are chemically
XX modified using a solution of bisulphite, hydrogen sulphite or disulphite.
XX Also disclosed are oligonucleotides and/or PNA oligomers for detecting
XX the cytosine methylation state (CpG islands) of these genes, and a method
XX for the diagnosis and/or therapy of genetic and epigenetic parameters of
XX genes associated with signal transduction. The genomic DNA can be
XX obtained from cells or cellular components which contain DNA, e.g. cell
XX lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,
XX tissue embedded in paraffin such as tissue from eyes, intestine, kidney,
XX brain, heart, prostate, lung, breast or liver, histologic object slides,
XX and all their possible combinations. The sequences of the invention are
XX useful for the diagnosis and therapy of diseases associated with signal
XX transduction e.g. solid tumours and cancer. ABK31158-ABK31545 represent
XX chemically pretreated genomic DNA sequences of different genes associated
XX with signal transduction, or their complementary sequences. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from the
XX European Patent Office
XX
XX Sequence 5956 BP; 1652 A; 171 C; 1433 G; 2700 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 6; Length 5956;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 287 TTTTATATATATAGATAT 307
Db 5381 TTTTATATATATAGATAT 5401

RESULT 5
ID ABL70326
ABL70326 standard; DNA; 5956 BP.
XX
XX ABL70326;
AC

ABK31369
ID ABK31369 standard; DNA; 5956 BP.
XX
AC ABK31369;
XX
DT 23-APR-2002 (first entry)
XX
DE Signal transduction associated gene modified complementary DNA #106.
XX
KW Human; signal transduction associated gene; cytosine methylation state;
KW CpG island; signal transduction associated disease; solid tumour; cancer;
KW antitumour; cytostatic; mutant; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX WO200200926-A2.
XX
XX 03-JAN-2002.
XX
XX 29-JUN-2001; 2001WO-EP007472.
XX
XX 30-JUN-2000; 2000DE-01032529.
XX
XX 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2002-147896/19.
XX
XX Oligonucleotide for diagnosis and therapy of diseases associated with
XX signal transduction e.g. cancer, comprises chemically modified genomic
XX sequences of genes associated with signal transduction.
XX
XX Claim 1; SEQ ID NO 212; 24pp; English.
XX
XX The present invention relates to chemically modified DNA sequences of
XX signal transduction associated genes. The DNA sequences are chemically
XX modified using a solution of bisulphite, hydrogen sulphite or disulphite.
XX Also disclosed are oligonucleotides and/or PNA oligomers for detecting
XX the cytosine methylation state (CpG islands) of these genes, and a method
XX for the diagnosis and/or therapy of genetic and epigenetic parameters of
XX genes associated with signal transduction. The genomic DNA can be
XX obtained from cells or cellular components which contain DNA, e.g. cell
XX lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,
XX tissue embedded in paraffin such as tissue from eyes, intestine, kidney,
XX brain, heart, prostate, lung, breast or liver, histologic object slides,
XX and all their possible combinations. The sequences of the invention are
XX useful for the diagnosis and therapy of diseases associated with signal
XX transduction e.g. solid tumours and cancer. ABK31158-ABK31545 represent
XX chemically pretreated genomic DNA sequences of different genes associated
XX with signal transduction, or their complementary sequences. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from the
XX European Patent Office
XX
XX Sequence 5956 BP; 1652 A; 171 C; 1433 G; 2700 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 6; Length 5956;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 287 TTTTATATATATAGATAT 307
Db 5381 TTTTATATATATAGATAT 5401

RESULT 6
AAS61273
ID AAS61273 standard; DNA; 5956 BP.
XX
AC AAS61273;
XX
XX 29-JAN-2002 (first entry)
XX
DE Human gene regulation-associated gene oligonucleotide #228.
XX
XX Human; Gene regulation-associated gene; severe combined immunodeficiency;
XX cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
XX asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
XX renal disease; Preeclampsia; cardiac allograft vascular disease;
```

KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
 KW immunostimulant; cardiant; antiinflammatory; coagulant; antiasthmatic;
 KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN WO200177375-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-EP003968.
 XX
 PR 06-APR-2000; 2000DE-01019058.
 PR 07-APR-2000; 2000DE-01019173.
 PR 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2002-017470/02.
 XX
 PT New nucleic acid sequences from chemically modified genes associated with
 PT gene regulation, useful for analyzing cytosine methylations for diagnosis
 PT and therapy of diseases e.g. severe combined immunodeficiency disease.
 XX
 PS Disclosure; SEQ ID NO 234; 26pp; English.
 XX
 CC The invention relates to 224 nucleic acid sequences comprising at least
 CC 18 bases of a chemically pretreated gene associated with gene regulation
 CC selected from 43 known genes (or complementary sequences). The chemical
 CC pretreatment converts cytosine bases unmethylated at the 5-position to
 CC uracil or another base with hybridisation behaviour dissimilar to
 CC cytosine, to enable analysis of cytosine methylations. The DNA sequences,
 CC oligomers (or sets/arrays) and method are useful in the diagnosis of
 CC diseases (or predisposition to diseases) associated with gene regulation
 CC and in therapy of such diseases, by enabling analysis of the cytosine
 CC methylation patterns of such genes. Kits are provided. They are
 CC especially useful in diagnosis and therapy of e.g. severe combined
 CC immunodeficiency disease, cardiac disorders, haemophilia, solid tumours
 CC and cancer, Werner syndrome, asthma, HDR syndrome, Saethre-Chotzen
 CC syndrome, renal disease, preclampsia, graft versus-host disease. The
 CC present sequence is a sequence included in the sequence data for this
 CC specification and is associated with the human gene regulation-associated
 CC genes. Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 5956 BP; 1652 A; 171 C; 1433 G; 2700 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 6; Length 5956;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 287 TTTTATATATATAGATAT 307
 DB 5381 TTTTATATATATAGATAT 5401

RESULT 7
 ABL07848
 ID ABL07848 standard; cDNA; 16739 BP.
 XX
 AC ABL07848;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 18026.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX

OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 XX
 DR P-PSDB; ABB63745.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Claim 1; SEQ ID NO 18026; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 16739 BP; 4986 A; 3249 C; 3289 G; 5215 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 4; Length 16739;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 209 TAAATTAATAATCTATCTAT 229
 DB 2143 TAAATTAATAATCTATCTAT 2163

RESULT 8
 ABL07850
 ID ABL07850 standard; cDNA; 21828 BP.
 XX
 AC ABL07850;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 18032.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.
 DR P-PSDB; ABB63747.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Claim 1; SEQ ID NO 18032; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 21828 BP; 6614 A; 4180 C; 4296 G; 6738 T; 0 U; 0 Other;
 Query Match 1.3%; Score 21; DB 4; Length 21828;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 209 TAAATTAAATTCATTCTAT 229
 Db |||||
 7232 TAAATTAAATTCATTCTAT 7252
 RESULT 9
 ABLN17338
 ID ABLN17338 standard; cDNA; 291 BP.
 AC ABLN17338;
 XX
 XX 24-JUN-2002 (first entry)
 XX
 DE Human ORFX polynucleotide sequence SEQ ID NO:3153.
 XX
 KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200192523-A2.
 XX
 XX 06-DEC-2001.
 XX
 XX 29-MAY-2001; 2001WO-US010836.
 XX
 XX 30-MAY-2000; 2000US-0206132P.
 XX 29-AUG-2000; 2000US-0228716P.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 XX Shimkets RA, Leach MD;
 XX WPI; 2002-106308/14.
 DR P-PSDB; ABP01586.
 XX
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders.
 XX

PS Disclosure; SEQ ID NO 3153; 1037pp; English.
 XX
 CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABLN15762 to ABLN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC syndrome in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis.
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 291 BP; 107 A; 41 C; 56 G; 86 T; 0 U; 1 Other;
 Query Match 1.2%; Score 20; DB 6; Length 291;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 68 TCATTTCAGAGGATGAAAAA 87
 Db |||||
 164 TCATTTCAGAGGATGAAAAA 183
 RESULT 10
 AAL15883/c
 ID AAL15883 standard; cDNA; 785 BP.
 XX
 AC AAL15883;
 XX
 DT 07-DEC-2001 (first entry)
 XX
 DE Human breast cancer expressed polynucleotide 8340.
 XX
 KW Human; breast cancer; cell marker; cytostatic; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200151628-A2.
 XX
 XX 19-JUL-2001.
 XX
 XX 10-JAN-2001; 2001WO-US000798.
 XX
 XX 14-JAN-2000; 2000US-0176077P.
 XX 14-MAR-2000; 2000US-0189167P.
 XX 24-MAR-2000; 2000US-0192099P.
 XX 29-MAR-2000; 2000US-0193480P.
 XX 15-MAY-2000; 2000US-0205230P.
 XX 09-JUN-2000; 2000US-0211315P.
 XX 25-JUL-2000; 2000US-0220534P.
 XX
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 XX Lillie J, Xu Y, Wang Y, Steinmann K;
 XX WPI; 2001-451856/48.
 XX
 XX New peptide useful as a marker for the diagnosis of breast cancer.
 PT

```

XX PS Claim 1; Page 1509; 3695pp; English.
XX CC The invention relates to human breast cancer expressed polynucleotides
XX CC (AAL07544-AAL26789) and methods of assessing whether a patient is
XX CC afflicted with breast cancer by examining the correlation between the
XX CC expression of certain markers and the cancerous state of breast cells.
XX CC The polynucleotides and encoded polypeptides are potential markers for
XX CC detecting, diagnosing, monitoring, characterising treating and
XX CC potentially preventing breast cancer. The polynucleotides and encoded
XX CC polypeptides are also useful for isolating compounds with cytostatic
XX CC activity
XX SQ Sequence 785 BP; 216 A; 172 C; 184 G; 200 T; 0 U; 13 Other;

Query Match 1.2%; Score 20; DB 4; Length 785;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 TTTTTCCTTAAGAAAAAAA 248
Db 709 TTTTTCCTTAAGAAAAAAA 690

RESULT 11
ADA49079
ID ADA49079 standard; DNA; 830 BP.
XX AC ADA49079;
XX DT 20-NOV-2003 (first entry)
XX DE Wheat gene conferring disease resistance in plants.
XX KW disease resistance; pathogen tolerance; plant pathogen; ds; gene; plant;
XX KW wheat.
XX OS Triticum aestivum.
XX PN WO2003000906-A2.
XX PD 03-JAN-2003.
XX PF 21-JUN-2002; 2002WO-IB002453.
XX PR 22-JUN-2001; 2001US-0300112P.
XX PR 26-SEP-2001; 2001US-0352277P.
XX PR 22-MAR-2002; 2002US-0366535P.
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX PI Glazebrook J, Briggs S, Cooper B, Goff SA, Moughamer T;
XX PI Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;
XX DR WPI; 2003-184052/18.
XX PT New polynucleotide comprising a plant nucleotide sequence having an open
XX PT reading frame that encodes a polypeptide associated with disease
XX PT resistance, useful for conferring resistance or tolerance to a plant
XX PT pathogen.
XX PS Disclosure; SEQ ID NO 1149; 299pp; English.
XX CC The invention relates to a novel isolated polynucleotide comprising a
XX CC plant nucleotide sequence having an open reading frame that encodes a
XX CC polypeptide associated with disease resistance or its fragment having
XX CC substantially the same activity as the full-length polypeptide. The
XX CC polynucleotide of the invention is useful for conferring resistance or
XX CC polynucleotide of the invention is useful for conferring resistance or
XX CC tolerance to a plant pathogen. The present sequence represents a gene
XX CC conferring disease resistance used in the invention.
XX SQ Sequence 830 BP; 177 A; 221 C; 219 G; 213 T; 0 U; 0 Other;

Query Match 1.2%; Score 20; DB 8; Length 830;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1613 GATCGCGCCGAGCTCGCCA 1632
Db 545 GATCGCGCCGAGCTCGCCA 564

RESULT 12
ACC61661
ID ACC61661 standard; DNA; 1044 BP.
XX AC ACC61661;
XX DT 20-JUN-2003 (first entry)
XX DE Gene sequence #SEQ ID 2104.
XX KW Multiprotein complex; eukaryote; drug target; diagnosis; gene; ds.
XX OS Saccharomyces cerevisiae.
XX PN BP1258494-A1.
XX PD 20-NOV-2002.
XX PF 20-DEC-2001; 2001EP-00130253.
XX PR 15-MAY-2001; 2001EP-00111774.
XX PA (CELL-) CELLZOME AG.
XX PI Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;
XX PI Marzioch M, Schultz JD, Superti-Furga GD;
XX DR WPI; 2003-250078/25.
XX DR P-PSDB; ABR53619.
XX PT New isolated protein complexes useful for diagnosing a disease or
XX PT disorder, or as a target for an active agent of a pharmaceutical,
XX PT preferably a drug target in the treatment or prevention of disease or
XX PT disorder.
XX PS Disclosure; SEQ ID NO 2104; 17pp + Sequence Listing; English.
XX CC The invention relates to multiprotein complexes from eukaryotes. Proteins
XX CC of the invention and DNA sequences encoding them are given in records
XX CC ABR52568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are
XX CC obtainable by using a protein as a bait and isolating the set of proteins
XX CC which is attached thereto from cells. Such protein complexes may comprise
XX CC up to 30 distinct proteins. Protein complexes of the invention are useful
XX CC for diagnosing a disease or disorder, or as a target for an active agent
XX CC of a pharmaceutical, preferably a drug target in the treatment or
XX CC prevention of a disease or disorder. Note: The sequence data for this
XX CC patent is not represented in the printed specification, but is based on
XX CC sequence information supplied by the European Patent Office. The complete
XX CC document is available on CD-ROM
XX SQ Sequence 1044 BP; 367 A; 164 C; 229 G; 284 T; 0 U; 0 Other;

Query Match 1.2%; Score 20; DB 7; Length 1044;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 TCATTTTCAGAGGATGAAAAA 87
Db 506 TCATTTTCAGAGGATGAAAAA 525

RESULT 13
AAZ52423/c
ID AAZ52423 standard; DNA; 1963 BP.

```

XX AAZ52423;
 XX AC
 XX DT 24-FEB-2000 (first entry)
 XX DE
 XX DE HTRM clone 2019742 DNA sequence.
 XX KW HTRM; human transcriptional regulatory molecule; arteriosclerosis; AIDS;
 KW arteriosclerosis; cirrhosis; cancer; leukaemia; diabetes mellitus; ss;
 KW Addison's disease; multiple sclerosis; rheumatoid arthritis; infection;
 KW trauma; myasthenia gravis; adenocarcinoma; immune disorder; treatment.
 XX OS Homo sapiens.
 XX PN W09957144-A2.
 XX XX 11-NOV-1999.
 XX PF 04-MAY-1999; 99WO-US009935.
 XX PR 05-MAY-1998; 98US-0084254P.
 XX PR 07-AUG-1998; 98US-0095827P.
 XX PR 02-OCT-1998; 98US-0102745P.
 XX PA (INCY-) INCYTE PHARM INC.
 XX PI Hillman JL, Bandman O, Lal P, Yue H, Reddy R, Tang YT;
 PI Gerstin EH, Patterson C, Baughn MR, Azimzai Y, Lu DAM;
 XX DR WPI; 2000-052941/04.
 XX DR P-PSDB; AAY73338.
 XX PT New peptides useful for diagnosis, prevention and treatment of cancer and
 PT immune disorders.
 XX XX Claim 9; Page 155; 193pp; English.
 XX XX AAZ52410-Z52474 are human transcriptional regulator molecule (HTRM)
 CC nucleotide sequences. The HTRM protein and nucleotide sequences are
 CC useful for preventing or treating disorders associated with decreased
 CC expression or activity of HTRM which include cell proliferative disorders
 CC such as arteriosclerosis and cirrhosis; cancers including adenocarcinoma
 CC and leukaemia; immune disorders such as AIDS, Addison's disease, diabetes
 CC mellitus, rheumatoid arthritis, multiple sclerosis, systemic lupus
 CC erythematosus, and myasthenia gravis; infections and trauma. Antagonists
 CC of the HTRM polypeptides are useful for treating or preventing disorders
 CC associated with increased expression or activity of HTRMs. HTRM
 CC screening libraries of immunogenic fragments or oligopeptides are useful for
 CC Polynucleotides encoding HTRM are useful for blocking the transcription
 CC of mRNA and regulating gene function by modulating the activity of HTRM.
 CC Vectors expressing HTRM or agonists can also be used to prevent or treat
 CC disorder associated with decreased HTRM expression. Antibodies which
 CC specifically bind HTRM and polynucleotides encoding HTRM are useful for
 CC diagnosing disorders associated with the expression of HTRM, particularly
 CC in assays that detect the expression of HTRM. Nucleotide sequences
 CC encoding HTRM may be useful to generate hybridization probes useful in
 CC mapping the naturally occurring genomic sequence and to detect
 CC differences in gene sequences among normal, carrier and affected
 CC individuals. Using diagnostic assays, cancer can be detected prior to the
 CC appearance of clinical symptoms and thereby progression of cancer can be
 CC prevented by aggressive treatment or preventive measures
 XX SQ Sequence 1963 BP; 488 A; 499 C; 491 G; 485 T; 0 U; 0 Other;
 Query Match 1.2%; Score 20; DB 3; Length 1963;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 236 TTAAGAAAAAACAGCCCA 255
 |||||
 1644 TTAAGAAAAAACAGCCCA 1625

RESULT 14
 AAI60732
 ID AAI60732 standard; cDNA; 1973 BP.
 XX AC
 XX AC AAI60732;
 XX DT 22-OCT-2001 (first entry)
 XX XX Human polynucleotide SEQ ID NO 4721.
 XX KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX OS Homo sapiens.
 XX PN W0200153312-A1.
 XX PD 26-JUL-2001.
 XX PF 26-DEC-2000; 2000WO-US034263.
 XX PR 23-DEC-1999; 99US-00471275.
 XX PR 21-JAN-2000; 2000US-00488725.
 XX PR 25-APR-2000; 2000US-00552317.
 XX PR 20-JUN-2000; 2000US-00598042.
 XX PR 19-JUL-2000; 2000US-00620312.
 XX PR 03-AUG-2000; 2000US-00653450.
 XX PR 14-SEP-2000; 2000US-00662191.
 XX PR 19-OCT-2000; 2000US-00693036.
 XX PR 29-NOV-2000; 2000US-00727344.
 XX PA (HVSE-) HYSEQ INC.
 XX XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX DR WPI; 2001-442253/47.
 XX DR P-PSDB; AAM41576.
 XX PT Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX PS Claim 1; SEQ ID NO 4721; 10078pp; English.
 XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with neotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX SQ Sequence 1973 BP; 485 A; 493 C; 505 G; 490 T; 0 U; 0 Other;
 Query Match 1.2%; Score 20; DB 4; Length 1973;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 236 TTAAGAAAAAACAGCCCA 255

Db 315 TTAAGAAAAAACAGCCCA 334
|||||

RESULT 15
AAI58946/c
ID AAI58946 standard; cDNA; 1979 BP.
XX AC AAI58946;
XX XX
DT 22-OCT-2001 (first entry)
XX XX
DE Human polynucleotide SEQ ID NO 1149.
XX XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX XX
XX Homo sapiens.
XX OS
XX WO200153312-A1.
XX PN
XX BD 26-JUL-2001.
XX XX
XX 26-DEC-2000; 2000WO-US034263.
XX XX
XX 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX XX
XX (HYSE-) HYSEQ INC.
XX XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX XX
XX WPI: 2001-442253/47.
DR P-PSDB; AAM39790.
XX XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX XX
XX Claim 1; SEQ ID NO 1149; 10078pp; English.
XX XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAI42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX XX
XX Query sequence 1979 BP; 498 A; 505 C; 485 G; 491 T; 0 U; 0 Other;
SQ

Query Match 1.2%; Score 20; DB 4; Length 1979;
Best Local Similarity 100.0%; Pred. No. 35;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 13:02:10 : Search time 144 Seconds
(without alignments)
6389.643 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattcaggctcaatac.....ctcgttagttggacggcg 1658

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A-COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B-COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A-COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B-COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PTCUS-COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	1.2	1979	4	US-09-620-312D-838
C 2	19	1.1	786431	4	US-09-751-389-3
C 3	19	1.1	1230025	4	US-09-198-452A-1
C 4	18	1.1	533	4	US-09-669-751-93
C 5	18	1.1	624	4	US-09-489-039A-4810
C 6	18	1.1	936	4	US-09-134-001C-1009
C 7	18	1.1	1428	4	US-09-328-352-1253
C 8	18	1.1	2311	2	US-08-712-709-6
C 9	18	1.1	2311	3	US-09-111-444-6
C 10	18	1.1	2311	3	US-09-541-228-6
C 11	18	1.1	2311	4	US-09-016-434-772
C 12	18	1.1	2370	4	US-09-031-295-1
C 13	18	1.1	3845	2	US-08-220-240A-4
C 14	18	1.1	3982	4	US-08-956-171E-520
C 15	18	1.1	4061	4	US-09-620-312D-363
C 16	18	1.1	5153	4	US-10-238-483-3
C 17	18	1.1	5658	4	US-09-647-344A-17
C 18	18	1.1	87350	3	US-08-781-891-79
C 19	18	1.1	87350	4	US-09-618-166-79
C 20	18	1.1	87543	4	US-09-791-211-3
C 21	18	1.1	4403765	3	US-09-103-840A-2
C 22	18	1.1	4411529	3	US-09-103-840A-1
C 23	17	1.0	243	1	US-08-341-568-2
C 24	17	1.0	243	2	US-08-911-020-2
C 25	17	1.0	285	4	US-09-134-001C-1300
C 26	17	1.0	291	4	US-09-184-418C-14
C 27	17	1.0	417	4	US-09-621-976-3318

C 28	17	1.0	498	4	US-09-252-991A-11412	Sequence 11412, A
C 29	17	1.0	621	4	US-09-107-532A-2883	Sequence 2883, Ap
C 30	17	1.0	666	1	US-08-463-115-4	Sequence 4, Appli
C 31	17	1.0	666	1	US-08-465-388-4	Sequence 4, Appli
C 32	17	1.0	714	4	US-09-540-236-757	Sequence 757, App
C 33	17	1.0	780	4	US-09-252-991A-3741	Sequence 3741, Ap
C 34	17	1.0	960	4	US-09-252-991A-3730	Sequence 3730, Ap
C 35	17	1.0	1071	4	US-09-252-991A-3633	Sequence 3633, Ap
C 36	17	1.0	1089	4	US-09-252-991A-3785	Sequence 3785, Ap
C 37	17	1.0	1209	3	US-09-105-537-21	Sequence 21, Appli
C 38	17	1.0	1278	4	US-09-252-991A-4932	Sequence 4932, Ap
C 39	17	1.0	1329	3	US-08-360-758-1	Sequence 1, Appli
C 40	17	1.0	1389	1	US-08-458-023B-1	Sequence 1, Appli
C 41	17	1.0	1389	3	US-09-111-556A-1	Sequence 1, Appli
C 42	17	1.0	1407	4	US-09-252-991A-11395	Sequence 11395, A
C 43	17	1.0	1594	2	US-08-955-713-1	Sequence 1, Appli
C 44	17	1.0	1717	4	US-09-634-137-3	Sequence 3, Appli
C 45	17	1.0	1857	3	US-09-299-378-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-620-312D-838/C
; Sequence 838, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yuning
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: Pt_FL_genes Version 1.0
; SEQ ID NO 838
; LENGTH: 1979
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (175)..(1053)
US-09-620-312D-838

Query Match 1.2%; Score 20; DB 4; Length 1979;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 236 TTAAGAAAAAACAAGCCCA 255
DB 1661 TTAAGAAAAAACAAGCCCA 1642


```
; LOCATION: (465001)..(480000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (480001)..(495000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (495001)..(510000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (510001)..(525000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (525001)..(540000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (540001)..(555000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (555001)..(570000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (570001)..(585000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (585001)..(600000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (600001)..(615000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (615001)..(630000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (630001)..(645000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (645001)..(660000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (660001)..(675000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (675001)..(690000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (690001)..(705000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (705001)..(720000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (720001)..(735000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (735001)..(750000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (750001)..(765000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (765001)..(780000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (780001)..(795000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (795001)..(810000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (810001)..(825000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (825001)..(840000)
```

```
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (840001)..(855000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (855001)..(870000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (870001)..(885000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (885001)..(900000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (900001)..(915000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
```

Query Match 1.1%; Score 19; DB 4; Length 1230025;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1039 GAGTGTAGAGAGAACTGA 1057
|||||

Db 1200765 GAGTGTAGAGAGAACTGA 1200783
|||||

RESULT 4

US-09-669-751-93
; Sequence 93, Application US/09669751
; Patent No. 6551575
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Ralph J.
; TITLE OF INVENTION: Methods for Identifying Compounds for
; TITLE OF INVENTION: Motion Sickness, Vertigo and Other Disorders Related to
; TITLE OF INVENTION: Balance and the Perception of Gravity
; FILE REFERENCE: P-NI 3864
; CURRENT APPLICATION NUMBER: US/09/669,751
; CURRENT FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 60/168,579
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Drosophila
US-09-669-751-93

Query Match 1.1%; Score 18; DB 4; Length 533;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 689 CTTTGTGTTTGTGTTTCT 706
|||||

Db 263 CTTTGTGTTTGTGTTTCT 280
|||||

RESULT 5

US-09-489-039A-4810
; Sequence 4810, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4810

```
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4810

Query Match          1.1%; Score 18; DB 4; Length 624;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1378 CGCGTCGCGCGCACGGG 1395
Db 394 CGCGTCGCGCGCACGGG 411

RESULT 6
US-09-134-001C-1009
; Sequence 1009, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1009
; LENGTH: 936
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1009

Query Match          1.1%; Score 18; DB 4; Length 936;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 CAAATATTTTAAATGAAT 120
Db 248 CAAATATTTTAAATGAAT 265

RESULT 7
US-09-328-352-1253
; Sequence 1253, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1253
; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1253

Query Match          1.1%; Score 18; DB 4; Length 1428;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 243 AAAAAACAGCCCATTAAG 260
Db 36 AAAAAACAGCCCATTAAG 53

RESULT 8
```

```
US-08-712-709-6/c
; Sequence 6, Application US/08712709
; Patent No. 5863780
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,709
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
US-08-712-709-6

Query Match          1.1%; Score 18; DB 2; Length 2311;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1501 CGGCTATAAAAGGCGGC 1518
Db 908 CGGCTATAAAAGGCGGC 891

RESULT 9
US-09-111-444-6/c
; Sequence 6, Application US/09111444
; Patent No. 6045792
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
```

```
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,444
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/712,709
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
; US-09-111-444-6

Query Match 1.1%; Score 18; DB 3; Length 2311;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1501 CGGCTATAAAAAAGCGGC 1518
Db 908 CGGCTATAAAAAAGCGGC 891

RESULT 10
US-09-541-228-6/c
; Sequence 6, Application US/09541228
; Patent No. 6232077
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/541,228
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/712,709
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
; US-09-541-228-6

Query Match 1.1%; Score 18; DB 3; Length 2311;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1501 CGGCTATAAAAAAGCGGC 1518
Db 908 CGGCTATAAAAAAGCGGC 891

RESULT 11
US-09-016-434-772/c
; Sequence 772, Application US/09016434
; Patent No. 650938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 772:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: MMLR2DT01
; CLONE: 477245
; US-09-016-434-772

Query Match 1.1%; Score 18; DB 4; Length 2311;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1501 CGGCTATAAAAAAGCGGC 1518
Db 908 CGGCTATAAAAAAGCGGC 891
```

```
RESULT 12
US-09-031-295-1/c
; Sequence 1, Application US/09031295
; Patent No. 6326181
; GENERAL INFORMATION:
; APPLICANT: LANG, Florian
; APPLICANT: WALDEGGER, Tübingen
; TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-SGK
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,295
; FILING DATE: 26-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 197-08-173.8
; FILING DATE: 28-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sandercock, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 058315/0123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2370 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 43..1335
US-09-031-295-1

Query Match 1.1%; Score 18; DB 4; Length 2370;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1501 CGGCTATAAAAGCGCGC 1518
Db 941 CGGCTATAAAAGCGCGC 924

RESULT 13
US-08-220-240A-4/c
; Sequence 4, Application US/08220240A
; Patent No. 5955291
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Matikainen, Marja-Terttu
; APPLICANT: Partanen, Juha
; APPLICANT: Makela, Tomi
; APPLICANT: Korkonen, Jaana
; TITLE OF INVENTION: ANTIBODIES RECOGNIZING TIE RECEPTOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
```

```
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/220,240A
FILING DATE: 29-MAR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FI93/00006
FILING DATE: 08-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/817,800
FILING DATE: 09-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/167,453
FILING DATE: 15-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 29151/31958
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3845 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 37..3450
US-08-220-240A-4

Query Match 1.1%; Score 18; DB 2; Length 3845;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 CACATAGGCGCTTCCTGGC 1171
Db 3390 CACATAGGCGCTTCCTGGC 3373

RESULT 14
US-08-956-171E-520
; Sequence 520, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
```

```
;
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/956,171E
;   FILING DATE: 20-Oct-1997
;   CLASSIFICATION: <unknown>
;
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 60/009,861
;   FILING DATE: January 5, 1996
;   APPLICATION NUMBER: 08/781,986
;   FILING DATE: January 3, 1997
;
; ATTORNEY/AGENT INFORMATION:
;   NAME: Mark J. Hyman
;   REGISTRATION NUMBER: 46,789
;   REFERENCE/DOCKET NUMBER: PB248P1
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (240) 314-1224
;     TELEFAX: (301) 309-8439
;
; INFORMATION FOR SEQ ID NO: 520:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 3982 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;
;   SEQUENCE DESCRIPTION: SEQ ID NO: 520:
;
; US-08-956-171E-520
;
; Query Match      1.1%; Score 18; DB 4; Length 3982;
; Best Local Similarity 100.0%; Pred. No. 35;
; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY      43 TTTATAGTTAGATTACTT 60
; DB      284 TTTATAGTTAGATTACTT 301
;
; RESULT 15
; US-09-620-312D-363
; Sequence 363, Application US/09620312D
; Patent NO. 659662
; GENERAL INFORMATION:
;   APPLICANT: Tang, Y. Tom
;   APPLICANT: Liu, Chenghua
;   APPLICANT: Asundi, Vinod
;   APPLICANT: Zhang, Jie
;   APPLICANT: Ren, Feiyan
;   APPLICANT: Chen, Rui-hong
;   APPLICANT: Zhao, Qing A.
;   APPLICANT: Wehrman, Tom
;   APPLICANT: Xue, Aidong J.
;   APPLICANT: Yang, Yonghong
;   APPLICANT: Wang, Jian-Rui
;   APPLICANT: Zhou, Ping
;   APPLICANT: Ma, Yungqing
;   APPLICANT: Wang, Dunrui
;   APPLICANT: Wang, Zhiwei
;   APPLICANT: John Tillinghast
;   APPLICANT: Drmanac, Radoje T.
;
; TITLE OF INVENTION: No. 659662el Nucleic Acids and
;   TITLE OF INVENTION: Polypeptides
;
; FILE REFERENCE: 784CIP2B
;
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
;
; NUMBER OF SEQ ID NOS: 1105
;
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 363
; LENGTH: 4061
; TYPE: DNA
;
;
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (72)..(3062)
; US-09-620-312D-363
;
; Query Match      1.1%; Score 18; DB 4; Length 4061;
; Best Local Similarity 100.0%; Pred. No. 35;
; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY      38 TAAATTTTATAGTTAGAT 55
; DB      3544 TAAATTTTATAGTTAGAT 3561
;
; Search completed: September 2, 2004, 21:26:12
; Job time : 152 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 14:41:35 ; Search time 783 Seconds
(without alignments)
10525.313 Million cell updates/sec

Title: US-10-732-721-1

Perfect score: 1658

Sequence: 1 gaattcaggctcaacaatac.....ctcgttagttgggacggcg 1658

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 3267054 seqs, 2485319735 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6534108

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1658	100.0	1658	17	US-10-732-721-1
2	67	4.0	836	13	Sequence 1, Appli
3	66	4.0	918	13	Sequence 7252, Ap
4	40	2.4	894	13	Sequence 13779, A
5	22	1.3	2509	17	Sequence 3295, Ap
6	21	1.3	21	17	Sequence 17878, A
7	21	1.3	21	17	Sequence 4, Appli
8	21	1.3	5956	13	Sequence 5, Appli
9	20	1.2	1089	15	Sequence 234, App
10	20	1.2	1291	13	Sequence 6995, Ap
11	20	1.2	1291	16	Sequence 151164,
12	20	1.2	1690	15	Sequence 151164,
13	20	1.2	1924	13	Sequence 13, Appl
14	20	1.2	1924	16	Sequence 204099,

c 15	20	1.2	1979	15	US-10-037-270-838	Sequence 838, App
c 16	20	1.2	1979	16	US-10-117-722-838	Sequence 838, App
c 17	20	1.2	3255	16	US-10-094-749-1102	Sequence 1102, Ap
c 18	20	1.2	5823	15	US-10-240-453-256	Sequence 256, App
c 19	20	1.2	16163	10	US-09-764-891-6505	Sequence 6505, Ap
c 20	20	1.2	16163	13	US-10-091-414-270	Sequence 270, App
c 21	20	1.2	65454	17	US-10-293-864-11	Sequence 11, Appl
c 22	20	1.2	100301	17	US-10-450-826-83	Sequence 83, Appl
c 23	20	1.2	3186778	13	US-10-027-632-174961	Sequence 174961,
c 24	20	1.2	3186778	16	US-10-027-632-174961	Sequence 174961,
c 25	19	1.1	399	16	US-10-369-493-29761	Sequence 29761, A
c 26	19	1.1	485	13	US-10-027-632-324809	Sequence 324809,
c 27	19	1.1	485	16	US-10-027-632-324809	Sequence 324809,
c 28	19	1.1	559	13	US-10-027-632-312341	Sequence 312341,
c 29	19	1.1	559	16	US-10-027-632-312341	Sequence 312341,
c 30	19	1.1	571	13	US-10-027-632-71076	Sequence 71076, A
c 31	19	1.1	571	16	US-10-027-632-71076	Sequence 71076, A
c 32	19	1.1	604	13	US-10-001-885-14	Sequence 14, Appl
c 33	19	1.1	613	13	US-10-027-632-288247	Sequence 288247,
c 34	19	1.1	613	13	US-10-027-632-288248	Sequence 288248,
c 35	19	1.1	613	16	US-10-027-632-288247	Sequence 288247,
c 36	19	1.1	613	16	US-10-027-632-288248	Sequence 288248,
c 37	19	1.1	619	13	US-10-027-632-266112	Sequence 266112,
c 38	19	1.1	619	13	US-10-027-632-266113	Sequence 266113,
c 39	19	1.1	619	16	US-10-027-632-266112	Sequence 266112,
c 40	19	1.1	619	16	US-10-027-632-266113	Sequence 266113,
c 41	19	1.1	640	13	US-10-027-632-64828	Sequence 64828, A
c 42	19	1.1	640	13	US-10-027-632-310725	Sequence 310725,
c 43	19	1.1	640	16	US-10-027-632-64828	Sequence 64828, A
c 44	19	1.1	640	16	US-10-027-632-310725	Sequence 310725,
c 45	19	1.1	732	13	US-10-027-632-103276	Sequence 103276,

ALIGNMENTS

RESULT 1

US-10-732-721-1
; Sequence 1, Application US/10732721
; Publication No. US20040163144A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
; FILE REFERENCE: 38-15(52826)A
; CURRENT APPLICATION NUMBER: US/10732,721
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/60/434,242
; PRIOR FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1658
; TYPE: DNA
; ORGANISM: Zea mays
US-10-732-721-1

Query Match 100.0%; Score 1658; DB 17; Length 1658;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1658; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GAATTCACGGCTCACAAATACAGTCATCTACATGTGATAAATTTATAGTTAGTACTT	60
Db	1	GAATTCACGGCTCACAAATACAGTCATCTACATGTGATAAATTTATAGTTAGTACTT	60
Qy	61	CTTGTAATCAATTTTACAGAGGATGAAAAAATCGCAGAAAGCAAAATTTTAATCAAT	120
Db	61	CTTGTAATCAATTTTACAGAGGATGAAAAAATCGCAGAAAGCAAAATTTTAATCAAT	120
Qy	121	GATGCAATATACAAATTTAATTTACAAATTTATGTAAGATTACATTTGTTAGTTTCATAGA	180
Db	121	GATGCAATATACAAATTTAATTTACAAATTTATGTAAGATTACATTTGTTAGTTTCATAGA	180

Qy	181	AATCAATTTCTAGAGTCATAAATAAGCCTAAATTAATAATCTATTCTATTTTTCCTTAAG	240
Db	181	AATCAATTTCTAGAGTCATAAATAAGCCTAAATTAATAATCTATTCTATTTTTCCTTAAG	240
Qy	241	AAAAAAACAGCCATTAAAGGGACCAATTAAGATGCGCTGTCTCCATTTTTTTATATATA	300
Db	241	AAAAAAACAGCCATTAAAGGGACCAATTAAGATGCGGGCTGCTCAATTTTTTTATATATA	300
Qy	301	GAGATATGAGTTGTGTCTACTAGGCCACCTCCACCTCGCANCGTGTTCTTTATTACATG	360
Db	301	GAGATATGAGTTGTGTCTACTAGGCCACCTCCACCTCGCATCGGTCTCTTTATTACATG	360
Qy	361	AAAAATGAGCCCTGCTTTTGCATTAAGGGTCTAGAGATGACGAGTGCACGCTTCCT	420
Db	361	AAAAATGAGCCCGTCTTTTGCATTAAGGGTCTAGAGATGACGAGTGCACGCTTCCT	420
Qy	421	ATTCACCTGCACGAGTAATTTTTTTTTTAATGACGAGGGCGCTTTGTATTCTCAGCAGCT	480
Db	421	ATTCACCTGCACGAGTAATTTTTTTTTTAATGACGAGGGCGCTTTGTATTCTCAGCAGCT	480
Qy	481	CTGGGCCCGTGTCTGCTGTA CGGCA GTGACAGAGCAGCAGCAGCTCTACGTGCCACG	540
Db	481	CTGGGCCCGTGTCTGCTGTA CGGCA GTGACAGAGCAGCAGCAGCTCTACGTGCCACG	540
Qy	541	TGCTGTGCTGCTGCCACCGTGATGCCAAGGCTTGCTTGTTCGTGCGTGCCTGCTGGC	600
Db	541	TGCTGTGCTGCTGCCACCGTGATGSCAAGGCTTGCTTGTTCGTGCGTGCCTGCTGGC	600
Qy	601	AGGACGAGGTTGAAGCTACGGCGGTGGCTGTGCTGCGCGCGTGTAGCACGGCGAC	660
Db	601	AGGACGAGGTTGAAGCTACGGCGGTGGCTGTGCTGCGCGCGTGTAGCACGGCGAC	660
Qy	661	GACCTCCAGCTGATGCTAGCGTTGCTACTTTTGTTTTGTCTCTGCTGCTTGGATG	720
Db	661	GACCTCAGCGTGA GTAGCGTTGCTACTTTTGTTTTGTCTCTGCTGCTTGGATG	720
Qy	721	ATGCAGTAAATAAACCCTGTGTATCTCATGTGATGTATCGACGCTGTGGCGATTCAA	780
Db	721	ATGCAGTAAATAAACCCTGTGTATCTCATGTGATGTATCGACGCTGTGGCGATTCAA	780
Qy	781	ATCAGTAAACATCATGCTGCATCCCATCGACTTCTGCTGCGCGATAAATCTGTAGGG	840
Db	781	ATCAGTAAACATCATGCTGCATCCCATCGACTTCTGCTGCGCGATAAATCTGTAGGG	840
Qy	841	ATTTGTGCTTTTCATGTATGTCTTGAATAGAGAGGAAGAATAATATACATAAAATAT	900
Db	841	ATTTGTGCTTTTCATGTATGTCTTGAATAGAGAGGAAGAATAATATACATAAAATAT	900
Qy	901	GAGTTCAACATCTTTCTTAATACCTTTTGTCTGATCTCTTTAGTAGAGCCAGTGGTGA	960
Db	901	GAGTTCAACATCTTTCTTAATACCTTTTGTCTGATCTCTTTAGTAGAGCCAGTGGTGA	960
Qy	961	TAACTGTTTAAAGAACCCCTGTATACCGAGTGTAGTCCAAAGCTGCTTAACTGAATAATAG	1020
Db	961	TAACTGTTTAAAGAACCCCTGTATACCGAGTGTAGTCCAAAGCTGCTTAACTGAATAATAG	1020
Qy	1021	TAGAACACAAGAGACAAGAGTGTAGAGAGGAACGTATCTTTGTATCTATATGTGCTG	1080
Db	1021	TAGAACACAAGAGACAAGAGTGTAGAGAGGAACGTATCTTTGTATCTATATGTGCTG	1080
Qy	1081	CTCTCCAAAGGTTACATGATATGGGATCTCTCTCTATTTATAGACAAACCTAGGGTTT	1140
Db	1081	CTCTCCAAAGGTTACATGATATGGGATCTCTCTCTATTTATAGACAAACCTAGGGTTT	1140
Qy	1141	CAGGCATATGGGCCACATAGGCCCTTCTTGCGCCCAAGAAAGGTTTCTTAAACCTACCT	1200
Db	1141	CAGGCATATGGGCCACATAGGCCCTTCTTGCGCCCAAGAAAGGTTTCTTAAACCTACCT	1200
Qy	1201	TCCAGGCGCGTGTGCTGCTTCACTCTCTTCCCTCGGCTCCAGAGCGGATAAATACG	1260
Db	1201	TCCAGGCGCGTGTGCTGCTTCACTCTCTTCCCTCGGCTCCAGAGCGGATAAATACG	1260
Qy	1261	GGGTCTACGGTGCACGCGCACGCGATGTGTGCGCGAGGGGCATCTCTCGCGGGGAATGCG	1320

Db	1261		GGGTCTACG	GTGAC	CGCGC	AGCAT	GTGCG	CCGAG	GGCAT	CTCTCC	CGGG	AATGCG	1320	
Qy	1321		CGCAGG	CCATCG	TGCAC	ACG	CAGT	GTGTG	CGCCCC	CGCAG	GGCTCT	CTCGC	ACAGC	1380
Db	1321		CGCAGG	CCATCG	TGCAC	ACG	CAGT	GTGTG	CGCCCC	CGCAG	GGCTCT	CTCGC	ACAGC	1380
Qy	1381		GTGCG	CGCGC	CAGG	GGCCAG	GGAC	CTTGCA	TTGTC	CGCGT	CGGTG	CACAC	GTCCCT	1440
Db	1381		GTGCG	CGCGC	CAGG	GGCCAG	GGAC	CTTGCA	TTGTC	CGCGT	CGGTG	CACAC	GTCCCT	1440
Qy	1441		GGTC	CGGCTT	GGCGG	CTT	GGCTT	GGCGG	CGAC	GTGCG	CGCTCT	CCCC	ACCGGA	1500
Db	1441		GGTC	CGGCTT	GGCGG	CTT	GGCTT	GGCGG	CGAC	GTGCG	CGCTCT	CCCC	ACCGGA	1500
Qy	1501		CGGCT	TAAAA	AGG	CGG	CGCAC	CTGAT	CTTC	CAC	AAGCA	AAGC	AGCAG	1560
Db	1501		CGGCT	TAAAA	AGG	CGG	CGCAC	CTGAT	CTTC	CAC	AAGCA	AAGC	AGCAG	1560
Qy	1561		AGCC	CAAC	CTTAA	ACTAA	AGAA	CAGT	AGT	AGT	CCCTGT	GAC	GTCTAG	1620
Db	1561		AGCC	CAAC	CTTAA	ACTAA	AGAA	CAGT	AGT	AGT	CCCTGT	GAC	GTCTAG	1620
Qy	1621		CCGAG	CTCG	CGC	AGT	TTT	GTCT	CGT	TAG	TTT	TGG	AGCGG	1680
Db	1621		CCGAG	CTCG	CGC	AGT	TTT	GTCT	CGT	TAG	TTT	TGG	AGCGG	1680

RESULT 2

```

US-10-425-114-7252
; Sequence 7252, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7252
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700623908_F1I
US-10-425-114-7252

```

Query Match 4.0%; Score 67; DB 13; Length 836;
Best Local Similarity 100.0%; Pred. No. 4.7e-24;
Matches 67: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1553	GOAGCAACAGCCCAACACCTAAACCTAAAGAACAGTAGTACTCCCTGTGTACGTTAGCAAC	1612
D _b	1	GCAGCAACAGCCCAACACCTAAACCTAAAGAACAGTAGTACTCCCTGTGTACGTTAGCAAC	60

Qy	1613	GATCGC	1619
Db	61	GATCGC	67

RESULT 3

RESULT 3
US-1C-425-114-13779
; Sequence 13779, Application US10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua

```
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425.114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 13779
; LENGTH: 918
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB143-042-D8_FLI
US-10-425-114-13779

Query Match          4.0%; Score 66; DB 13; Length 918;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1554 CAGCAACAGGCAACACCTAACTAAAGAACAGTAGTAGTCCCTGTGTACGTTAGCAACG 1613
      |||||||
Db 1 CAGCAACAGGCAACACCTAACTAAAGAACAGTAGTAGTCCCTGTGTACGTTAGCAACG 60

QY 1614 ATCCGC 1619
      ||||||
Db 61 ATCCGC 66

RESULT 4
US-10-425-114-3295
; Sequence 3295, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425.114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3295
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700258323_FLI
US-10-425-114-3295

Query Match          2.4%; Score 40; DB 13; Length 894;
Best Local Similarity 100.0%; Pred. No. 7.2e-10;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1580 GAACAGTAGTAGTCCCTGTGTACGTTAGCAACGATCCGC 1619
      |||||||
Db 1 GAACAGTAGTAGTCCCTGTGTACGTTAGCAACGATCCGC 40

RESULT 5
US-10-437-963-17878/c
; Sequence 17878, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```

```
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437.963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 17878
; LENGTH: 2509
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_23489C.1
US-10-437-963-17878

Query Match          1.3%; Score 22; DB 17; Length 2509;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 657 CGACGAGCTCGACGCTGATGCT 678
      |||||||
Db 568 CGACGAGCTCGACGCTGATGCT 547

RESULT 6
US-10-732-721-4
; Sequence 4, Application US/10732721
; Publication No. US20040163144A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
; FILE REFERENCE: 38-15(52826)A
; CURRENT APPLICATION NUMBER: US/10/732.721
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/60/434,242
; PRIOR FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: PAT_MRT4530_23489C.1
US-10-732-721-4

Query Match          1.3%; Score 21; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1226 ACTCCTTCCTCCGCTCCAG 1246
      |||||||
Db 1 ACTCCTTCCTCCGCTCCAG 21

RESULT 7
US-10-732-721-5/c
; Sequence 5, Application US/10732721
; Publication No. US20040163144A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
; FILE REFERENCE: 38-15(52826)A
; CURRENT APPLICATION NUMBER: US/10/732.721
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/60/434,242
; PRIOR FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
```

; SEQ ID NO 5
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Zea mays
US-10-732-721-5

Query Match 1.3%; Score 21; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1638 GCTCGTTAGTTGGAGCGCG 1658
|||||
Db 21 GCTCGTTAGTTGGAGCGCG 1

RESULT 8

US-10-221-613-234
; Sequence 234, Application US/10221613
; Publication No. US20040029123A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221,613
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 234
; LENGTH: 5956
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-234

Query Match 1.3%; Score 21; DB 13; Length 5956;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 287 TTTTATATATAGATAT 307
|||||
Db 5381 TTTTATATATAGATAT 5401

RESULT 9

US-10-198-846-6995/c
; Sequence 6995, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220

; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6995
; LENGTH: 1089
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5, 6, 7, 10, 11, 12, 13, 16, 17, 18, 21, 22, 23, 28, 29,
; LOCATION: 30, 31, 33, 34, 35, 39, 41, 47, 49, 50, 53, 54, 56, 58,
; LOCATION: 59, 60, 61, 62, 63, 65, 66, 67, 69, 73, 74, 75, 79, 80,
; LOCATION: 82, 83, 84, 85, 86, 87, 88, 89, 92, 93, 94, 97, 99, 103,
; LOCATION: 107
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 109, 110, 111, 112, 113, 114, 115, 116, 124, 182, 202, 203,
; LOCATION: 204, 228, 406, 831, 870, 914, 921, 927, 929, 932, 947, 954,
; LOCATION: 972, 977, 980, 983, 994, 995, 997, 1017, 1032, 1043, 1050,
; LOCATION: 1053, 1081, 1083, 1085, 1088
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-6995

Query Match 1.2%; Score 20; DB 15; Length 1089;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 TTTTCTTAAAGAAAAA 248
|||||
Db 867 TTTTCTTAAAGAAAAA 848

RESULT 10

US-10-027-632-151164
; Sequence 151164, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151164
; LENGTH: 1291
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-151164

Query Match 1.2%; Score 20; DB 13; Length 1291;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1578 AAGACAGTAGTACCTG 1597
|||||
Db 276 AAGACAGTAGTACCTG 295

RESULT 11
US-10-027-632-151164
; Sequence 151164, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151164
; LENGTH: 1291
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-151164

Query Match 1.2%; Score 20; DB 16; Length 1291;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1578 AAGAACAGTAGTAGTCCTG 1597
|||||
Db 276 AAGAACAGTAGTAGTCCTG 295

RESULT 12
US-10-268-822-13/c
; Sequence 13, Application US/10268822
; Publication No. US20030150004A1
; GENERAL INFORMATION:
; APPLICANT: Moore, David
; APPLICANT: Wei, Ping
; APPLICANT: Chua, Steven
; TITLE OF INVENTION: Screening Systems and Methods for Identifying Modulators of Xenob
; TITLE OF INVENTION: Metabolism
; FILE REFERENCE: P02729US2
; CURRENT APPLICATION NUMBER: US/10/268,822
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: PCT/US 01/29672
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 10/219,590
; PRIOR FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 1690
; TYPE: DNA
; ORGANISM: Mouse
US-10-268-822-13

Query Match 1.2%; Score 20; DB 15; Length 1690;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 900 TGAGTTCAACATCTTCTTA 919
|||||
Db 1227 TGAGTTCAACATCTTCTTA 1208

RESULT 13
US-10-027-632-204099/c
; Sequence 204099, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 204099
; LENGTH: 1924
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1924)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-204099

Query Match 1.2%; Score 20; DB 13; Length 1924;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 TTTCCTTAAGAAAAAACA 250
|||||
Db 1770 TTTCCTTAAGAAAAAACA 1751

RESULT 14
US-10-027-632-204099/c
; Sequence 204099, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 204099
; LENGTH: 1924
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1924)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-204099

Query Match 1.2%; Score 20; DB 16; Length 1924;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 TTTCCTTAAGAAAAAACA 250
|||
Db 1770 TTTCCTTAAGAAAAAACA 1751

RESULT 15

US-10-037-270-838/c
; Sequence 838, Application US/10037270
; Publication NO. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Felyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1 Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 838
; LENGTH: 1979
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (175)..(1053)
US-10-037-270-838

Query Match 1.2%; Score 20; DB 15; Length 1979;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 TTAAGAAAAAACAGCCCA 255
|||
Db 1661 TTAAGAAAAAACAGCCCA 1642

Search completed: September 2, 2004, 21:49:23
Job time : 788 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 12:48:50 ; Search time 4953 Seconds
(without alignments)
11778.434 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattcacggctcacataac.....ctcgtagttgggacggcg 1658

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 37577330 seqs, 17593059518 residues

Word size : 0

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending Patents NA Main:*

- 1: /cgn2_6/ptodata/2/pna/PCTUS COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06 COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07 COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US08 COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US082 COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US083 COMB.seq.*
- 8: /cgn2_6/ptodata/2/pna/US084 COMB.seq.*
- 9: /cgn2_6/ptodata/2/pna/US085 COMB.seq.*
- 10: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
- 11: /cgn2_6/ptodata/2/pna/US087 COMB.seq.*
- 12: /cgn2_6/ptodata/2/pna/US088 COMB.seq.*
- 13: /cgn2_6/ptodata/2/pna/US089 COMB.seq.*
- 14: /cgn2_6/ptodata/2/pna/US090 COMB.seq.*
- 15: /cgn2_6/ptodata/2/pna/US091 COMB.seq.*
- 16: /cgn2_6/ptodata/2/pna/US092 COMB.seq.*
- 17: /cgn2_6/ptodata/2/pna/US093 COMB.seq.*
- 18: /cgn2_6/ptodata/2/pna/US094 COMB.seq.*
- 19: /cgn2_6/ptodata/2/pna/US095 COMB.seq.*
- 20: /cgn2_6/ptodata/2/pna/US096 COMB.seq.*
- 21: /cgn2_6/ptodata/2/pna/US097 COMB.seq.*
- 22: /cgn2_6/ptodata/2/pna/US098 COMB.seq.*
- 23: /cgn2_6/ptodata/2/pna/US099 COMB.seq.*
- 24: /cgn2_6/ptodata/2/pna/US099A COMB.seq.*
- 25: /cgn2_6/ptodata/2/pna/US099B COMB.seq.*
- 26: /cgn2_6/ptodata/2/pna/US099C COMB.seq.*
- 27: /cgn2_6/ptodata/2/pna/US099D COMB.seq.*
- 28: /cgn2_6/ptodata/2/pna/US099E COMB.seq.*
- 29: /cgn2_6/ptodata/2/pna/US099F COMB.seq.*
- 30: /cgn2_6/ptodata/2/pna/US099G COMB.seq.*
- 31: /cgn2_6/ptodata/2/pna/US099H COMB.seq.*
- 32: /cgn2_6/ptodata/2/pna/US099I COMB.seq.*
- 33: /cgn2_6/ptodata/2/pna/US099J COMB.seq.*
- 34: /cgn2_6/ptodata/2/pna/US099K COMB.seq.*
- 35: /cgn2_6/ptodata/2/pna/US099L COMB.seq.*
- 36: /cgn2_6/ptodata/2/pna/US099M COMB.seq.*
- 37: /cgn2_6/ptodata/2/pna/US099N COMB.seq.*
- 38: /cgn2_6/ptodata/2/pna/US099O COMB.seq.*
- 39: /cgn2_6/ptodata/2/pna/US099P COMB.seq.*
- 40: /cgn2_6/ptodata/2/pna/US099Q COMB.seq.*
- 41: /cgn2_6/ptodata/2/pna/US099R COMB.seq.*
- 42: /cgn2_6/ptodata/2/pna/US099S COMB.seq.*
- 43: /cgn2_6/ptodata/2/pna/US100A COMB.seq.*

- 44: /cgn2_6/ptodata/2/pna/US100B COMB.seq.*
- 45: /cgn2_6/ptodata/2/pna/US101A COMB.seq.*
- 46: /cgn2_6/ptodata/2/pna/US101B COMB.seq.*
- 47: /cgn2_6/ptodata/2/pna/US102A COMB.seq.*
- 48: /cgn2_6/ptodata/2/pna/US102B COMB.seq.*
- 49: /cgn2_6/ptodata/2/pna/US103A COMB.seq.*
- 50: /cgn2_6/ptodata/2/pna/US103B COMB.seq.*
- 51: /cgn2_6/ptodata/2/pna/US104A COMB.seq.*
- 52: /cgn2_6/ptodata/2/pna/US104B COMB.seq.*
- 53: /cgn2_6/ptodata/2/pna/US106 COMB.seq.*
- 54: /cgn2_6/ptodata/2/pna/US107A COMB.seq.*
- 55: /cgn2_6/ptodata/2/pna/US107B COMB.seq.*
- 56: /cgn2_6/ptodata/2/pna/US6000 COMB.seq.*
- 57: /cgn2_6/ptodata/2/pna/US6001 COMB.seq.*
- 58: /cgn2_6/ptodata/2/pna/US6002 COMB.seq.*
- 59: /cgn2_6/ptodata/2/pna/US6003 COMB.seq.*
- 60: /cgn2_6/ptodata/2/pna/US6004 COMB.seq.*
- 61: /cgn2_6/ptodata/2/pna/US6005 COMB.seq.*
- 62: /cgn2_6/ptodata/2/pna/US6006 COMB.seq.*
- 63: /cgn2_6/ptodata/2/pna/US6007 COMB.seq.*
- 64: /cgn2_6/ptodata/2/pna/US6008 COMB.seq.*
- 65: /cgn2_6/ptodata/2/pna/US6009 COMB.seq.*
- 66: /cgn2_6/ptodata/2/pna/US6010 COMB.seq.*
- 67: /cgn2_6/ptodata/2/pna/US6011 COMB.seq.*
- 68: /cgn2_6/ptodata/2/pna/US6012 COMB.seq.*
- 69: /cgn2_6/ptodata/2/pna/US6013 COMB.seq.*
- 70: /cgn2_6/ptodata/2/pna/US6014 COMB.seq.*
- 71: /cgn2_6/ptodata/2/pna/US6015 COMB.seq.*
- 72: /cgn2_6/ptodata/2/pna/US6016 COMB.seq.*
- 73: /cgn2_6/ptodata/2/pna/US6017 COMB.seq.*
- 74: /cgn2_6/ptodata/2/pna/US6018 COMB.seq.*
- 75: /cgn2_6/ptodata/2/pna/US6019 COMB.seq.*
- 76: /cgn2_6/ptodata/2/pna/US6020 COMB.seq.*
- 77: /cgn2_6/ptodata/2/pna/US6021 COMB.seq.*
- 78: /cgn2_6/ptodata/2/pna/US6022 COMB.seq.*
- 79: /cgn2_6/ptodata/2/pna/US6023A COMB.seq.*
- 80: /cgn2_6/ptodata/2/pna/US6023B COMB.seq.*
- 81: /cgn2_6/ptodata/2/pna/US6024 COMB.seq.*
- 82: /cgn2_6/ptodata/2/pna/US6025 COMB.seq.*
- 83: /cgn2_6/ptodata/2/pna/US6026 COMB.seq.*
- 84: /cgn2_6/ptodata/2/pna/US6027 COMB.seq.*
- 85: /cgn2_6/ptodata/2/pna/US6028 COMB.seq.*
- 86: /cgn2_6/ptodata/2/pna/US6029 COMB.seq.*
- 87: /cgn2_6/ptodata/2/pna/US6030 COMB.seq.*
- 88: /cgn2_6/ptodata/2/pna/US6031 COMB.seq.*
- 89: /cgn2_6/ptodata/2/pna/US6032 COMB.seq.*
- 90: /cgn2_6/ptodata/2/pna/US6033 COMB.seq.*
- 91: /cgn2_6/ptodata/2/pna/US6034 COMB.seq.*
- 92: /cgn2_6/ptodata/2/pna/US6035 COMB.seq.*
- 93: /cgn2_6/ptodata/2/pna/US6036 COMB.seq.*
- 94: /cgn2_6/ptodata/2/pna/US6037 COMB.seq.*
- 95: /cgn2_6/ptodata/2/pna/US6038 COMB.seq.*
- 96: /cgn2_6/ptodata/2/pna/US6039 COMB.seq.*
- 97: /cgn2_6/ptodata/2/pna/US6040 COMB.seq.*
- 98: /cgn2_6/ptodata/2/pna/US6041 COMB.seq.*
- 99: /cgn2_6/ptodata/2/pna/US6042 COMB.seq.*
- 100: /cgn2_6/ptodata/2/pna/US6043 COMB.seq.*
- 101: /cgn2_6/ptodata/2/pna/US6044 COMB.seq.*
- 102: /cgn2_6/ptodata/2/pna/US6045 COMB.seq.*
- 103: /cgn2_6/ptodata/2/pna/US6046 COMB.seq.*
- 104: /cgn2_6/ptodata/2/pna/US6047 COMB.seq.*
- 105: /cgn2_6/ptodata/2/pna/US6048 COMB.seq.*
- 106: /cgn2_6/ptodata/2/pna/US6049 COMB.seq.*
- 107: /cgn2_6/ptodata/2/pna/US6050 COMB.seq.*
- 108: /cgn2_6/ptodata/2/pna/US6051 COMB.seq.*
- 109: /cgn2_6/ptodata/2/pna/US6052 COMB.seq.*
- 110: /cgn2_6/ptodata/2/pna/US6053 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1658	100.0	1658	100	US-60-434-242-1
2	1658	100.0	1658	110	US-60-531-039-16
3	79	4.8	385	26	US-09-620-111B-7324
4	68	4.1	308	18	US-09-304-517A-40406
5	68	4.1	308	19	US-09-371-146A-40406
6	68	4.1	308	42	US-09-985-678-40406
7	67	4.0	396	35	US-09-865-439A-30469
8	67	4.0	396	76	US-60-207-458-74766
9	67	4.0	470	35	US-09-865-439A-27134
10	67	4.0	470	76	US-60-207-458-71431
11	67	4.0	836	46	US-10-155-881-37309
12	67	4.0	836	51	US-10-425-114-7252
13	67	4.0	836	51	US-10-425-114A-7252
14	67	4.0	836	88	US-60-312-544-973
15	66	4.0	861	27	US-09-654-617-273716
16	66	4.0	861	29	US-09-684-016-273716
17	66	4.0	918	51	US-10-425-114-13779
18	66	4.0	918	51	US-10-425-114A-13779
19	64	3.9	204	18	US-09-304-517A-39577
20	64	3.9	204	19	US-09-371-146A-39577
21	64	3.9	204	42	US-09-985-678-39577
22	60	3.6	527	35	US-09-865-439A-48132
23	60	3.6	527	76	US-60-207-458-92395
24	59	3.6	280	18	US-09-303-031A-3994
25	59	3.6	280	19	US-09-304-517A-57027
26	59	3.6	280	19	US-09-371-146A-57027
27	59	3.6	280	36	US-09-894-949-3994
28	59	3.6	280	36	US-09-894-949A-3994
29	59	3.6	280	42	US-09-985-678-57027
30	52	3.1	437	18	US-09-304-517A-79013
31	52	3.1	437	19	US-09-371-146A-79013
32	52	3.1	437	24	US-09-565-306-1774
33	52	3.1	437	42	US-09-985-678-79013
34	50	3.0	460	35	US-09-865-439A-43255
35	50	3.0	460	76	US-60-207-458-87528
36	47	2.8	686	35	US-09-865-439A-31938
37	47	2.8	686	76	US-60-207-458-76235
38	43	2.6	315	18	US-09-304-517A-78600
39	43	2.6	315	19	US-09-371-146A-78600
40	43	2.6	315	24	US-09-565-306-3420
41	43	2.6	315	42	US-09-985-678-78600
42	42	2.5	316	35	US-09-865-439A-44537
43	42	2.5	316	76	US-60-207-458-88800
44	40	2.4	291	18	US-09-304-517A-39265
45	40	2.4	291	19	US-09-371-146A-39265

ALIGNMENTS

RESULT 1
US-60-434-242-1
; Sequence 1, Application US/60434242
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 38-15(52826)A
; CURRENT APPLICATION NUMBER: US/60/434, 242
; CURRENT FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1658
; TYPE: DNA
; ORGANISM: Zea mays
US-60-434-242-1
Query Match 100.0%; Score 1658; DB 100; Length 1658;


```
; NUMBER OF SEQ ID NOS: 295529
; SEQ ID NO 40406
; LENGTH: 308
; TYPE: DNA
; ORGANISM: Zea mays
US-09-985-678-40406

Query Match
Best Local Similarity 4.1%; Score 68; DB 42; Length 308;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1552 AGCAGCAACAGCCCAACACCTAACTAAAGAACAGTAGTCCCTGTGTAGCTCTAGCAA 1611
Db 1 AGCAGCAACAGCCCAACACCTAACTAAAGAACAGTAGTCCCTGTGTAGCTCTAGCAA 60

QY 1612 CGATCCGC 1619
Db 61 CGATCCGC 68

RESULT 7
US-09-865-439A-30469
; Sequence 30469, Application US/09865439A
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51936)B
; CURRENT APPLICATION NUMBER: US/09/865,439A
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,458
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 119126
; SEQ ID NO 30469
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3354-052-P1-K1-G5
US-09-865-439A-30469

Query Match
Best Local Similarity 4.0%; Score 67; DB 35; Length 396;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCCAACACCTAACTAAAGAACAGTAGTCCCTGTGTAGCTCTAGCAA 1612
Db 1 GCAGCAACAGCCCAACACCTAACTAAAGAACAGTAGTCCCTGTGTAGCTCTAGCAA 60

QY 1613 GATCCGC 1619
Db 61 GATCCGC 67

RESULT 8
US-60-207-458-74766
; Sequence 74766, Application US/60207458
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; APPLICANT: Conner, Timothy W.
; APPLICANT: Deikman, Jill
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ruan, Yijun G.
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Sammons, R. Douglas
; APPLICANT: Shukla, Hridayabhiranjan
; APPLICANT: Wu, Kunsheng
; APPLICANT: Xu, Nanfei
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
```

```
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 38-21(51936)A
; CURRENT APPLICATION NUMBER: US/60/207,458
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 152403
; SEQ ID NO 74766
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: LIB3354-052-P1-K1-G5
US-60-207-458-74766

Query Match
Best Local Similarity 4.0%; Score 67; DB 76; Length 396;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCCAACACCTAACTAAAGAACAGTAGTCCCTGTGTAGCTCTAGCAA 1612
Db 1 GCAGCAACAGCCCAACACCTAACTAAAGAACAGTAGTCCCTGTGTAGCTCTAGCAA 60

QY 1613 GATCCGC 1619
Db 61 GATCCGC 67

RESULT 9
US-09-865-439A-27134
; Sequence 27134, Application US/09865439A
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51936)B
; CURRENT APPLICATION NUMBER: US/09/865,439A
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,458
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 119126
; SEQ ID NO 27134
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(470)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3354-008-P1-K1-B11
US-09-865-439A-27134

Query Match
Best Local Similarity 4.0%; Score 67; DB 35; Length 470;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCCAACACCTAACTAAAGAACAGTAGTCCCTGTGTAGCTCTAGCAA 1612
Db 1 GCAGCAACAGCCCAACACCTAACTAAAGAACAGTAGTCCCTGTGTAGCTCTAGCAA 60

QY 1613 GATCCGC 1619
Db 61 GATCCGC 67

RESULT 10
US-60-207-458-71431
; Sequence 71431, Application US/60207458
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; APPLICANT: Conner, Timothy W.
; APPLICANT: Deikman, Jill
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
```

; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ruan, Yijun G.
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Sammons, R. Douglas
; APPLICANT: Shukla, Hridayabhiranjan
; APPLICANT: Wu, Kunsheng
; APPLICANT: Xu, Nanfei
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(51936)A
; CURRENT APPLICATION NUMBER: US/60/207,458
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 152403
; SEQ ID NO 71431
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: LIB3354-008-P1-K1-B11
US-10-732-721-1

Query Match 4.0%; Score 67; DB 76; Length 470;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1553 GCAGCAACAGCCAAACACCTAAAGAACAGTAGTAGTCCTCGTGTACGCTTAGCAAC 1612
Db 1 GCAGCAACAGCCAAACACCTAAAGAACAGTAGTAGTCCTCGTGTACGCTTAGCAAC 60
QY 1613 GATCCGC 1619
Db 61 GATCCGC 67

RESULT 11
US-10-155-881-37309
; Sequence 37309, Application US/10155881
; GENERAL INFORMATION:
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Lutfiyva, Linda L.
; APPLICANT: McIninch, James
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(15300)J
; CURRENT APPLICATION NUMBER: US/10/155,881
; CURRENT FILING DATE: 2002-05-22
; NUMBER OF SEQ ID NOS: 37595
; SEQ ID NO 37309
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
US-10-155-881-37309

Query Match 4.0%; Score 67; DB 46; Length 836;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1553 GCAGCAACAGCCAAACACCTAAAGAACAGTAGTAGTCCTCGTGTACGCTTAGCAAC 1612
Db 1 GCAGCAACAGCCAAACACCTAAAGAACAGTAGTAGTCCTCGTGTACGCTTAGCAAC 60
QY 1613 GATCCGC 1619
Db 61 GATCCGC 67

RESULT 12
US-10-425-114-7252
; Sequence 7252, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7252
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700623908_FLI
US-10-425-114-7252

Query Match 4.0%; Score 67; DB 51; Length 836;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1553 GCAGCAACAGCCAAACACCTAAAGAACAGTAGTAGTCCTCGTGTACGCTTAGCAAC 1612
Db 1 GCAGCAACAGCCAAACACCTAAAGAACAGTAGTAGTCCTCGTGTACGCTTAGCAAC 60
QY 1613 GATCCGC 1619
Db 61 GATCCGC 67

RESULT 13
US-10-425-114A-7252
; Sequence 7252, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7252
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700623908_FLI
US-10-425-114A-7252

Query Match 4.0%; Score 67; DB 51; Length 836;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1553 GCAGCAACAGCCAAACACCTAAAGAACAGTAGTAGTCCTCGTGTACGCTTAGCAAC 1612
Db 1 GCAGCAACAGCCAAACACCTAAAGAACAGTAGTAGTCCTCGTGTACGCTTAGCAAC 60
QY 1613 GATCCGC 1619
Db 61 GATCCGC 67

RESULT 14
US-60-312-544-973
; Sequence 973, Application US/60312544
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei

; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)A
; CURRENT APPLICATION NUMBER: US/60/312,544
; CURRENT FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 10730
; SEQ ID NO 973
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (108)..(449)
; OTHER INFORMATION: Clone ID: 700623908_FLI
US-60-312-544-973

Query Match 4.0%; Score 67; DB 88; Length 836;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCACACCTAACTAAAGAACAGTAGTAGTCCCTGCTGACGTCTAGCAAC 1612
Db 1 GCAGCAACAGCCACACCTAACTAAAGAACAGTAGTAGTCCCTGCTGACGTCTAGCAAC 60

QY 1613 GATCCGC 1619
Db 61 GATCCGC 67

RESULT 15
US-09-654-617-273716
; Sequence 273716, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 273716
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Zea mays
US-09-654-617-273716

Query Match 4.0%; Score 66; DB 27; Length 861;
Best Local Similarity 100.0%; Pred. No. 4.4e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1554 CAGCAACAGCCACACCTAACTAAAGAACAGTAGTAGTCCCTGCTGACGTCTAGCAACG 1613
Db 24 CAGCAACAGCCACACCTAACTAAAGAACAGTAGTAGTCCCTGCTGACGTCTAGCAACG 83

QY 1614 ATCCGC 1619
Db 84 ATCCGC 89

Search completed: September 2, 2004, 21:23:30
Job time : 4955 secs

This Page Blank (uspto)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1658	100.0	1658	7	US-10-732-721-1	Sequence 1, Appli
2	69	4.2	975	6	US-10-425-115-181022	Sequence 181022,
3	40	2.4	1034	6	US-10-425-115-181033	Sequence 181033,
C 4	37	2.3	528	6	US-10-425-115-47987	Sequence 47987, A
C 5	38	2.2	2211	6	US-10-425-115-82158	Sequence 82158, A
C 6	23	1.4	868	6	US-10-425-115-68623	Sequence 68623, A
7	21	1.3	21	7	US-10-732-721-4	Sequence 4, Appli
C 8	21	1.3	21	7	US-10-732-721-5	Sequence 5, Appli
C 9	21	1.3	772	6	US-10-425-115-30821	Sequence 30821, A
C 10	21	1.3	8344	6	US-10-918-754-16910	Sequence 16910, A
C 11	21	1.3	24289	5	US-10-918-754-16866	Sequence 16866, A
C 12	20	1.2	2891	5	US-09-865-590A-3153	Sequence 3153, Ap
13	20	1.2	751	6	US-10-425-115-12790	Sequence 12790, A
C 14	20	1.2	955	6	US-10-425-115-163648	Sequence 163648,
C 15	20	1.2	1690	1	PCT-US03-32273-13	Sequence 13, Appl
C 16	20	1.2	1973	1	PCT-US04-07412-1190	Sequence 1190, Ap
C 17	20	1.2	1973	8	US-10-389-559-1190	Sequence 1190, Ap
C 18	20	1.2	12988	6	US-10-472-965-729	Sequence 729, App
C 19	20	1.2	12988	6	US-10-895-164-8392	Sequence 8392, Ap
C 20	20	1.2	12988	6	US-10-896-164-11776	Sequence 11776, A
C 21	20	1.2	12988	8	US-10-868-184A-8392	Sequence 8392, Ap
C 22	20	1.2	12988	8	US-10-868-184A-11776	Sequence 11776, A
C 23	20	1.2	12988	8	US-10-868-184A-8392	Sequence 8392, Ap
C 24	20	1.2	12988	8	US-10-868-184-11776	Sequence 11776, A


```
Query Match          2.4%; Score 40; DB 6; Length 1034;
Best Local Similarity 100.0%; Pred. No. 5.9e-10;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1580 GAACAGTAGTACGCTCGTGTACGCTAGCAACGATCCGC 1619
      |||||||
Db 1 GAACAGTAGTACGCTCGTGTACGCTAGCAACGATCCGC 40

RESULT 4
US-10-425-115-47987/c
; Sequence 47987, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 47987
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(528)
; OTHER INFORMATION: unsure at all n locations
;
; OTHER INFORMATION: Clone ID: MPT4577_14376C.1
US-10-425-115-47987

Query Match          2.3%; Score 38; DB 6; Length 528;
Best Local Similarity 100.0%; Pred. No. 6.4e-09;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 944 TAGAGCCAGTCGCTGATACGCTGTAAAGAACCCCTTG 981
      |||||||
Db 109 TAGAGCCAGTCGCTGATACGCTGTAAAGAACCCCTTG 72

RESULT 5
US-10-425-115-82158/c
; Sequence 82158, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 82158
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
;
; OTHER INFORMATION: Clone ID: MPT4577_174942C.1
US-10-425-115-82158

Query Match          2.2%; Score 37; DB 6; Length 2211;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1088 AAGTTTACATGATAGGGATCTCCTCTATTATTATA 1124
      |||||||
```

```
Db 2209 AAGTTTACATGATAGGGATCTCCTCTATTATTATA 2173

RESULT 6
US-10-425-115-68623/c
; Sequence 68623, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 68623
; LENGTH: 868
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
;
; OTHER INFORMATION: Clone ID: MPT4577_162578C.1
US-10-425-115-68623

Query Match          1.4%; Score 23; DB 6; Length 868;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AGTCATCTACATGTGATAAATTT 44
      |||||||
Db 535 AGTCATCTACATGTGATAAATTT 513

RESULT 7
US-10-732-721-4
; Sequence 4, Application US/10732721
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
; FILE REFERENCE: 38-15(52826)A
; CURRENT APPLICATION NUMBER: US/10/732,721
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/60/434,242
; PRIOR FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Zea mays
US-10-732-721-4

Query Match          1.3%; Score 21; DB 7; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1226 ACTCCTTCCCTCCGCTCCAG 1246
      |||||||
Db 1 ACTCCTTCCCTCCGCTCCAG 21

RESULT 8
US-10-732-721-5/c
; Sequence 5, Application US/10732721
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
; FILE REFERENCE: 38-15(52826)A
; CURRENT APPLICATION NUMBER: US/10/732,721
; CURRENT FILING DATE: 2003-12-10
```

PRIOR APPLICATION NUMBER: US/60/434,242
PRIOR FILING DATE: 2002-12-18
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.2
SEQ ID NO 5
LENGTH: 21
TYPE: DNA
ORGANISM: Zea mays
US-10-732-721-5

Query Match 1.3%; Score 21; DB 7; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1638 GCTCGTACTTTGGGACGGCG 1658
DB 21 GCTCGTACTTTGGGACGGCG 1

RESULT 9
US-10-425-115-30821
Sequence 30821, Application US/10425115
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 30821
LENGTH: 772
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_128119C.1
US-10-425-115-30821

Query Match 1.3%; Score 21; DB 6; Length 772;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 ATTTCCTTAAGAAAAA 248
DB 480 ATTTCCTTAAGAAAAA 500

RESULT 10
US-10-754-16910/c
Sequence 16910, Application US/10918754
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: CL001480
CURRENT APPLICATION NUMBER: US/10/918,754
CURRENT FILING DATE: 2004-08-16
NUMBER OF SEQ ID NOS: 91238
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16910
LENGTH: 88344
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(88344)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-918-754-16910

Query Match 1.3%; Score 21; DB 6; Length 88344;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 AAGAAAGCAAAATATTTTAAAT 116
DB 47515 AAGAAAGCAAAATATTTTAAAT 47495

RESULT 11
US-10-918-754-16866/c
Sequence 16866, Application US/10918754
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: CL001480
CURRENT APPLICATION NUMBER: US/10/918,754
CURRENT FILING DATE: 2004-08-16
NUMBER OF SEQ ID NOS: 91238
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16866
LENGTH: 244289
TYPE: DNA
ORGANISM: Homo sapiens
US-10-918-754-16866

Query Match 1.3%; Score 21; DB 6; Length 244289;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 ABAATCTATTCTATTTC 235
DB 77392 ABAATCTATTCTATTTC 77372

RESULT 12
US-09-865-590A-3153
Sequence 3153, Application US/09865590A
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.
APPLICANT: Shimkets, Richard
TITLE OF INVENTION: Novel Human Polynucleotides and Polypeptides Encoded Thereby
FILE REFERENCE: 21402-016
CURRENT APPLICATION NUMBER: US/09/865,590A
CURRENT FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: 60/206,132
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: 60/228,716
PRIOR FILING DATE: 2000-08-29
NUMBER OF SEQ ID NOS: 22982
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3153
LENGTH: 291
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (274)..(274)
OTHER INFORMATION: Wherein n may be a, c, g or t
US-09-865-590A-3153

Query Match 1.2%; Score 20; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 TCATTTCAGAGATCAAAA 87
DB 164 TCATTTCAGAGATCAAAA 183

RESULT 13

PRIOR APPLICATION NUMBER: US/60/434,242
PRIOR FILING DATE: 2002-12-18
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.2
SEQ ID NO 5
LENGTH: 21
TYPE: DNA
ORGANISM: Zea mays
US-10-732-721-5

Query Match 1.3%; Score 21; DB 7; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1638 GCTCGTACTTTGGGACGGCG 1658
DB 21 GCTCGTACTTTGGGACGGCG 1

RESULT 9
US-10-425-115-30821
Sequence 30821, Application US/10425115
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 30821
LENGTH: 772
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_128119C.1
US-10-425-115-30821

Query Match 1.3%; Score 21; DB 6; Length 772;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 ATTTCCTTAAGAAAAA 248
DB 480 ATTTCCTTAAGAAAAA 500

RESULT 10
US-10-754-16910/c
Sequence 16910, Application US/10918754
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: CL001480
CURRENT APPLICATION NUMBER: US/10/918,754
CURRENT FILING DATE: 2004-08-16
NUMBER OF SEQ ID NOS: 91238
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16910
LENGTH: 88344
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(88344)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-918-754-16910

Query Match 1.3%; Score 21; DB 6; Length 88344;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 AAGAAAGCAAAATATTTTAAAT 116
DB 47515 AAGAAAGCAAAATATTTTAAAT 47495

RESULT 11
US-10-918-754-16866/c
Sequence 16866, Application US/10918754
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: CL001480
CURRENT APPLICATION NUMBER: US/10/918,754
CURRENT FILING DATE: 2004-08-16
NUMBER OF SEQ ID NOS: 91238
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16866
LENGTH: 244289
TYPE: DNA
ORGANISM: Homo sapiens
US-10-918-754-16866

Query Match 1.3%; Score 21; DB 6; Length 244289;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 ABAATCTATTCTATTTC 235
DB 77392 ABAATCTATTCTATTTC 77372

RESULT 12
US-09-865-590A-3153
Sequence 3153, Application US/09865590A
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.
APPLICANT: Shimkets, Richard
TITLE OF INVENTION: Novel Human Polynucleotides and Polypeptides Encoded Thereby
FILE REFERENCE: 21402-016
CURRENT APPLICATION NUMBER: US/09/865,590A
CURRENT FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: 60/206,132
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: 60/228,716
PRIOR FILING DATE: 2000-08-29
NUMBER OF SEQ ID NOS: 22982
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3153
LENGTH: 291
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (274)..(274)
OTHER INFORMATION: Wherein n may be a, c, g or t
US-09-865-590A-3153

Query Match 1.2%; Score 20; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 TCATTTCAGAGATCAAAA 87
DB 164 TCATTTCAGAGATCAAAA 183

RESULT 13

```
US-10-425-115-12790
; Sequence 12790, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 12790
; LENGTH: 751
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_111661C.1
US-10-425-115-12790

Query Match          1.2%; Score 20; DB 6; Length 751;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 958 TGATAACGCTGTTAAGAACCC 977
      |||||
Db 723 TGATAACGCTGTTAAGAACCC 742

RESULT 14
US-10-425-115-163648
; Sequence 163648, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 163648
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_80823C.1
US-10-425-115-163648

Query Match          1.2%; Score 20; DB 6; Length 955;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 958 TGATAACGCTGTTAAGAACCC 977
      |||||
Db 883 TGATAACGCTGTTAAGAACCC 902

RESULT 15
PCT-US03-32273-13/c
; Sequence 13, Application PC/TUS0332273
; GENERAL INFORMATION:
; APPLICANT: Moore, David
; APPLICANT: Wei, Ping
; APPLICANT: Chua, Steven
; TITLE OF INVENTION: Screening Systems and Methods for Identifying Modulators of Xenob
; FILE REFERENCE: P02729W02
; CURRENT APPLICATION NUMBER: PCT/US03/32273
; CURRENT FILING DATE: 2003-10-09
```

```
; PRIOR APPLICATION NUMBER: PCT/US 01/29672
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 10/268,822
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 10/219,590
; PRIOR FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 1690
; TYPE: DNA
; ORGANISM: Mouse
PCT-US03-32273-13

Query Match          1.2%; Score 20; DB 1; Length 1690;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 900 TGAGTTCAACATCTTTCTTA 919
      |||||
Db 1227 TGAGTTCAACATCTTTCTTA 1208

Search completed: September 2, 2004, 21:36:05
Job time : 588 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 10:46:24 ; Search time 3928 seconds
(without alignments)
12604.760 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattacggtcacaatac.....ctcgttagttggagcggcg 1658

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estmd:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_htc:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_htc:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	em_gss_hum:*
18:	em_gss_inv:*
19:	em_gss_pln:*
20:	em_gss_vrt:*
21:	em_gss_fun:*
22:	em_gss_mam:*
23:	em_gss_mus:*
24:	em_gss_pro:*
25:	em_gss_rod:*
26:	em_gss_phg:*
27:	em_gss_vrl:*
28:	gb_gss1:*
29:	gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	451	27.2	779	28	CC418961
C 2	436	26.3	728	29	CG152000
C 3	374	22.6	498	29	CG151999
C 4	313	18.9	754	29	CG146202

5	305	18.4	938	29	CG146205
6	281	16.9	373	29	CC654144
7	248	15.0	949	28	BZ705097
C 8	86	5.2	728	29	CG164746
C 9	67	4.0	786	29	CG062076
C 10	67	4.0	858	29	CG062075
C 11	64	3.9	438	29	CG193723
C 12	64	3.9	446	29	CG193725
C 13	64	3.9	732	28	CC352433
C 14	64	3.9	884	28	CC006284
C 15	60	3.6	844	29	CG096594
C 16	60	3.6	914	29	CG124360
C 17	60	3.6	925	29	CG096596
C 18	59	3.6	870	11	AY106704
C 19	56	3.4	879	29	CG118361
C 20	55	3.3	740	29	CG041830
C 21	48	2.9	656	28	BH872036
C 22	48	2.9	775	28	BZ961663
C 23	48	2.9	826	28	BZ988278
C 24	46	2.8	862	29	CG175932
C 25	46	2.8	985	29	CG175933
C 26	44	2.7	866	29	CG118362
C 27	38	2.3	528	10	AW216120
C 28	38	2.3	794	28	BZ991151
C 29	37	2.2	574	10	AW155643
C 30	37	2.2	665	28	CC443358
C 31	34	2.1	754	28	BZ828297
C 32	34	2.1	776	29	CG113123
C 33	34	2.1	802	28	CC382778
C 34	34	2.1	807	28	BZ828291
C 35	34	2.1	949	29	CG041829
C 36	34	2.1	993	29	CC463184
C 37	30	1.8	225	28	BZ750783
C 38	30	1.8	955	29	CG097258
C 39	29	1.7	656	29	CG050059
C 40	29	1.7	656	29	CG050060
C 41	29	1.7	695	29	CC752611
C 42	28	1.7	856	29	CG266334
C 43	28	1.7	910	29	CC615699
C 44	28	1.7	949	29	CG266345
C 45	26	1.6	1032	29	CG046479

ALIGNMENTS

RESULT 1
CC418961/c
LOCUS
DEFINITION
CC418961
Genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished (2003)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

CC418961
PUEDX81TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBtra230M18,
779 bp DNA linear GSS 19-MAY-2003
Genomic survey sequence.

CC418961
GSS
GI:30899051

CC418961.1
Zea mays
Zea mays

Unpublished (2003)
Contact: Cathy Whitelaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

```

FEATURES
source
  Location/Qualifiers
    1..779
      /organism="Zea mays"
      /mol_type="genomic DNA"
      /strain="B73"
      /db_xref="taxon:4577"
      /clone="ZMMBTa230M18"
      /clone_lib="ZM_0.6_1.0_KB"
      /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
      Cor selected genomic DNA library"

ORIGIN
Query Match      27.2%; Score 451; DB 28; Length 779;
Best Local Similarity 99.8%; Pred. No. 4.6e-203;
Matches 501; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 338 GCATCGGTCTCTTTATTACATGAAAAATGACCGCTCTCTTTTGCAATTAAGGGCTAGA 397
DB 502 GCATCGGTCTCTTTATTACATGAAAAATGACCGCTCTCTTTTGCAATTAAGGGCTAGA 443

QY 398 GATGACGAGTGAACGCTTCGTATTCATCGACGAGTAATTTTTTTTAATGACGGAG 457
DB 442 GATGACGAGTGAACGCTTCGTATTCATCGACGAGTAATTTTTTTTAATGACGGAG 383

QY 458 GCGCTTTGATTCCTCAGCAGCTTCGCGCGCCGCTTCGTGTCAGCGCAGTGACGAGC 517
DB 382 GCGCTTTGATTCCTCAGCAGCTTCGCGCGCCGCTTCGTGTCAGCGCAGTGACGAGC 323

QY 518 AGCAGCAGCTCTACGTGCCCCAGTGCCTGCTGCTGCCACCGCTGATGGCAGGCTTCG 577
DB 322 AGCAGCAGCTCTACGTGCCCCAGTGCCTGCTGCCACCGCTGATGGCAGGCTTCG 263

QY 578 TTGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 637
DB 262 TTGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 203

QY 638 TCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 697
DB 202 TCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 143

QY 698 TTGTTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 757
DB 142 TTGTTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 83

QY 758 TCGACGTGCTGCGCGATTTCAATCAGTAAACATCAGTGCATCCCGCTGCTGCTGCTG 817
DB 82 TCGACGTGCTGCGCGATTTCAATCAGTAAACATCAGTGCATCCCGCTGCTGCTGCTGCTG 23

QY 818 TGTGCGCGATAAATCTGCTAGG 839
DB 22 TGTGCGCGATAAATCTGCTAGG 1

RESULT 2
CG152000/c
LOCUS
DEFINITION
  728 bp DNA linear GSS 21-AUG-2003
  PUIFT17TD_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMMBTa0573C10,
  genomic survey sequence.
ACCESSION
  CG152000
VERSION
  1
KEYWORDS
  GSS.
SOURCE
  Zea mays
  Zea mays
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 728)
REFERENCE
  Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
  Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J., and
  Bennetzen, J.
  Maize Genomics Consortium
  Unpublished (2003)
  Other_GSSs: PUIFT17TB

TITLE
  JOURNAL
  COMMENT
  CG151999 498 bp DNA linear GSS 21-AUG-2003
  PUIFT17TB_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMMBTa0573C10,
  DEFINITION

```

```

Contact: Cathy Whitelaw
TIGR
712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seg primer: TP
Class: sheared ends.
Location/Qualifiers
  1..728
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMMBTa0573C10"
    /clone_lib="ZM_0.6_1.0_KB"
    /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
    Cor selected genomic DNA library"

```

ORIGIN

```

Query Match      26.3%; Score 436; DB 29; Length 728;
Best Local Similarity 99.5%; Pred. No. 6.4e-196;
Matches 656; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 601 AGCGACGAGTTGAAGCTACGCGGTGGCTGTGCTTGCCTCGCGCGCTGCTAGCACGCGGAC 660
DB 728 AGCGACGAGTTGAAGCTACGCGGTGGCTGTGCTTGCCTCGCGCGCTGCTAGCACGCGGAC 669

QY 661 GACGTCGACGTGATGCTAGCGTTGCTACTTTTGTGTTTTTGTGTTTTCTCTGCTTGGGATG 720
DB 668 GACGTCGACGTGATGCTAGCGTTGCTACTTTTGTGTTTTTGTGTTTTCTCTGCTTGGGATG 609

QY 721 ATGCGATTAATAAAGCTGCTGATCTCATGTCATGTCATGTCGTCGTCGCGGATTTCAA 780
DB 608 ATGCGATTAATAAAGCTGCTGATCTCATGTCATGTCATGTCGTCGTCGCGGATTTCAA 549

QY 781 ATCAGTAAACAATCACATGTCATCCCATCGACTTCGCTCGCGCGATAAATCTGCTAGG 840
DB 548 ATCAGTAAACAATCACATGTCATCCCATCGACTTCGCTCGCGCGATAAATCTGCTAGG 489

QY 841 ATTGTGCTTTTCATGTATGCTGCTGAATAGAGAGAAAGATAATATACATAAATAAT 900
DB 488 ATTGTGCTTTTCATGTATGCTGCTGAATAGAGAGAAAGATAATATACATAAATAAT 429

QY 901 GAGTTCACATCTTTCTTAATACCTTTTGTCTCGATCTCTCTCTCTATTATAGACAAAACAGGTT 960
DB 428 GAGTTCACATCTTTCTTAATACCTTTTGTCTCGATCTCTCTCTCTATTATAGACAAAACAGGTT 369

QY 961 TAACGTGTTAAGAACCCCTTGTGTACCGAGTGTAGTCCAA-GGCTGCTTAACGTGAATAA 1019
DB 368 TAACGTGTTAAGAACCCCTTGTGTACCGAGTGTAGTCCAAAGGGCTGCTTAACGTGAATAA 309

QY 1020 GTAGAACACAGAGACAGAGTGTAGAGAGAACTGATCTTTGTACTATATGCTGCT 1079
DB 308 GTAGAACACAGAGACAGAGTGTAGAGAGAACTGATCTTTGTACTATATGCTGCT 249

QY 1080 GCTCTCCAAAGGTTACATGATATGGGATCTCTCTCTCTATTATAGACAAAACAGGTT 1139
DB 248 GCTCTCCAAAGGTTACATGATATGGGATCTCTCTCTCTATTATAGACAAAACAGGTT 189

QY 1140 TCAGGCATATGGGCCACATAGGCTTTCCTGGCCCAAGAAAGGTTTCTTAACATCAATC 1199
DB 188 TCAGGCATATGGGCCACATAGGCTTTCCTGGCCCAAGAAAGGTTTCTTAACATCAATC 129

QY 1200 TTCAGGCGGCTGCTGCTGCTTCAACTCTTCCCTCCGCTCCAGAGCGGATAAATA 1258
DB 128 TTCAGGCGGCTGCTGCTGCTTCAACTCTTCCCTCCGCTCCAGAGCGGATAAATA 70

```

RESULT 3

```

CG151999
LOCUS
DEFINITION
  CG151999 498 bp DNA linear GSS 21-AUG-2003
  PUIFT17TB_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMMBTa0573C10,

```

```

genomic survey sequence.
ACCESSION CG151999
VERSION CG151999.1 GI:34042782
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 498)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PU1F7L7TD
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
Location/Qualifiers
source 1..498
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA0573C10"
/clone_lib="ZM 0.6 1.0 kb"
/note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"
ORIGIN
Query Match 22.6%; Score 374; DB 29; Length 498;
Best Local Similarity 99.8%; Pred. No. 2.2e-166;
Matches 424; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 338 GCATCGGTGTTCTTTATTTACATGAAAAAATGAGCGCTGCTTTTGCAATTAAGGGTCTAGA 397
Db 74 GCATCGGTGTTCTTTATTTACATGAAAAAATGAGCGCTGCTTTTGCAATTAAGGGCTAGA 133
QY 398 GATGTACAGTGCAACGCTTCGTATCTACTGACAGTAATTTTTTTTATGACCGAG 457
Db 134 GATGTACAGTGCAACGCTTCGTATCTACTGACAGTAATTTTTTTTATGACCGAG 193
QY 458 GCGCTTTGTATTCCTCAGCAGCTCTGCGCGCCGCTTCGTGTACGGCAGTGACGAGGC 517
Db 194 GCGCTTTGTATTCCTCAGCAGCTCTGCGCGCCGCTTCGTGTACGGCAGTGACGAGGC 253
QY 518 AGCAGCAGCTTACGTGCCACGCTGCTGCTGCTGCCACCGTGATGGAAGGCTTGC 577
Db 254 AGCAGCAGCTTACGTGCCACGCTGCTGCTGCTGCCACCGTGATGGAAGGCTTGC 313
QY 578 TTGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 637
Db 314 TTGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 373
QY 638 TGGCGCGTGTAGCAGCGGACGACGCTGACGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 697
Db 374 TGGCGCGTGTAGCAGCGGACGACGCTGACGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 433
QY 698 TTGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 757
Db 434 TTGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 493
QY 758 TCGAC 762
Db 494 TCGAC 498

```

```

RESULT 4
CG146202/c
LOCUS CG146202 754 bp DNA linear GSS 21-AUG-2003
DEFINITION PUJCV14TB ZM 0.6 1.0 kb Zea mays genomic clone ZMMBTA0639C04,
genomic survey sequence.
ACCESSION CG146202
VERSION CG146202.1 GI:34036985
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 754)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PUJCV14TD
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
Location/Qualifiers
source 1..754
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA0639C04"
/clone_lib="ZM 0.6 1.0 kb"
/note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"
ORIGIN
Query Match 18.9%; Score 313; DB 29; Length 754;
Best Local Similarity 99.7%; Pred. No. 1.9e-137;
Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1252 ATAAATACGGGGTCTACGGTGCACGCGCAGCATGGTGCAGGGGCGCATCTCTGCCG 1311
Db 364 ATAAATACGGGGTCTACGGTGCACGCGCAGCATGGTGCAGGGGCGCATCTCTGCCG 305
QY 1312 GGGATGCGCGCAGCGCATGCTCGACGACGACGACGACGACGACGACGACGACGACGAC 1371
Db 304 GGGATGCGCGCAGCGCATGCTCGACGACGACGACGACGACGACGACGACGACGACGAC 245
QY 1372 CGCACACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1431
Db 244 CGCACACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 185
QY 1432 CAGTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1491
Db 184 CAGTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 125
QY 1492 CCCACCGGCGGCTATATAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1551
Db 124 CCCACCGGCGGCTATATAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 65
QY 1552 AGCAGCAACAGCCAAACACCTAACTAAAGAACAGTAGTAGTCCCTGTGTACGTCTAGCAA 1611
Db 64 AGCAGCAACAGCCAAACACCTAACTAAAGAACAGTAGTAGTCCCTGTGTACGTCTAGCAA 5
QY 1612 CGAT 1615
Db 4 CGAT 1

```

RESULT 5
 CG146205 938 bp DNA linear GSS 21-AUG-2003
 LOCUS PUCJV147D ZM 0.6 1.0 KB Zea mays genomic clone ZMMBTa0639C04,
 DEFINITION genomic survey sequence.
 ACCESSION CG146205
 VERSION CG146205.1 GI:34036988
 KEYWORDS GSS.
 SOURCE CG146205.1
 ORGANISM Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 938)
 White14w,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
 Benwick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
 Bennettzen,J.
 Maize Genomics Consortium
 Unpublished (2003)
 Other_GSSs: PUCJV147B
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TP
 Class: sheared ends.
 Location/Qualifiers
 1..938
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBTa0639C04"
 /clone_lib="ZM 0.6 1.0 KB"
 /notes="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
 Cof selected genomic DNA library"
 ORIGIN
 Query Match 18.4%; Score 305; DB 29; Length 938;
 Best Local Similarity 99.7%; Pred. No. 1.1e-133;
 Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 338 GCATGCGTGTCTTTATTACATGAAAAAATGACGCGCTGCTTTTGCATTAAAGGCTCTAGA 397
 Db 563 GCATGCGTGTCTTTATTACATGAAAAAATGACGCGCTGCTTTTGCATTAAAGGCGCTAGA 622
 QY 398 GATGTACGAGTGCACACGCTTCGTATTTCACATGCACGAGTAATTTTTTTTAAATGCACGAG 457
 Db 623 GATGTACGAGTGCACACGCTTCGTATTTCACATGCACGAGTAATTTTTTTTAAATGCACGAG 682
 QY 458 GCGCTTTGTATTCTCTACGAGCTCTGCGCGCCCGCTGCTTCGTGTACGCGCAGTGCACGAGGC 517
 Db 683 GCGCTTTGTATTCTCTACGAGCTCTGCGCGCCCGCTGCTTCGTGTACGCGCAGTGCACGAGGC 742
 QY 518 AGCAGCAGCTCTACGTGCCACACGTCGCTGGCTGCTGCCACCGCTGATGCGCAAGGCTTGC 577
 Db 743 AGCAGCAGCTCTACGTGCCACACGTCGCTGGCTGCTGCCACCGCTGATGCGCAAGGCTTGC 802
 QY 578 TTGCTTCGCTGCTCGGTGGCAGCAGAGTTGAAGCTACGCGCGTGGCTGCTCTGCTTGC 637
 Db 803 TTGCTTCGCTGCTCGGTGGCAGCAGAGTTGAAGCTACGCGCGTGGCTGCTCTGCTTGC 862
 QY 638 TGC CGCGGTGCTAGCACGGCGACACGTCGACGCTGATGCTAGGCTGCTACTTTT 693
 Db 863 TGC CGCGGTGCTAGCACGGCGACACGTCGACGCTGATGCTAGGCTGCTACTTTT 918


```

ORGANISM      Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 949)
AUTHORS      Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE        Maize Genomics Consortium
JOURNAL      Unpublished (2003)
COMMENT      Contact: Cathy Whitelaw
TIGR         9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES
source       1..949
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA127G13"
/clone_lib="ZM 0.6-1.0 KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"
ORIGIN
Query Match      15.0%; Score 248; DB 28; Length 949;
Best Local Similarity 100.0%; Pred. No. 1.4e-106;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAATTCACGGCTCAATACACAGTCATCTACATGTGATAAATTTTATAGTTAGATTACTT 60
Db 409 GAATTCACGGCTCAATACACAGTCATCTACATGTGATAAATTTTATAGTTAGATTACTT 350
QY 61 CTTCGTATCATTTTCAGAGGTGAAAAAATCCAGAAAGCAAAATTTTAAATGAAT 120
Db 349 CTTCGTATCATTTTCAGAGGTGAAAAAATCCAGAAAGCAAAATTTTAAATGAAT 290
QY 121 GATGCAATATACAAATTTAATTCACAAATATGTAAGATTACATGTTTGTTCATAGA 180
Db 289 GATGCAATATACAAATTTAATTCACAAATATGTAAGATTACATGTTTGTTCATAGA 230
QY 181 AATCAATTTCTAGAGTCATAAATGCTAAATTTAAATTTCTATTTTCTTTAAG 240
Db 229 AATCAATTTCTAGAGTCATAAATGCTAAATTTAAATTTCTATTTTCTTTAAG 170
QY 241 AAAAAAAA 248
Db 169 AAAAAAAA 162
RESULT 8
CG164746/c
LOCUS
DEFINITION    CG164746 728 bp DNA linear GSS 21-AUG-2003
PUII051TB ZM 0.6-1.0 KB Zea mays genomic clone ZMMBTA0593105,
genomic survey sequence.
ACCESSION     CG164746
VERSION
KEYWORDS      GSS.
SOURCE
ORGANISM      Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 728)
AUTHORS      Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE        Maize Genomics Consortium
JOURNAL      Unpublished (2003)
COMMENT      Contact: Cathy Whitelaw
TIGR         9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES
source       1..728
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA0593105"
/clone_lib="ZM 0.6-1.0 KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"
ORIGIN
Query Match      5.2%; Score 86; DB 29; Length 728;
Best Local Similarity 100.0%; Pred. No. 1.7e-29;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1000 GGCTGCTTAAGTGAATTAATAGTACACACAGGAGACAAAGAGTGTAGAGAGAACTGATT 1059
Db 669 GGCTGCTTAAGTGAATTAATAGTACACACAGGAGACAAAGAGTGTAGAGAGAACTGATT 610
QY 1060 CTTTGTCTTACTATATGGTGTCTCTC 1085
Db 609 CTTTGTCTTACTATATGGTGTCTCTC 584
RESULT 9
CG062076/c
LOCUS
DEFINITION    CG062076 786 bp DNA linear GSS 19-AUG-2003
PUPPD96TD ZM 0.6-1.0 KB Zea mays genomic clone ZMMBTA0701024,
genomic survey sequence.
ACCESSION     CG062076
VERSION
KEYWORDS      GSS.
SOURCE
ORGANISM      Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 786)
AUTHORS      Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE        Maize Genomics Consortium
JOURNAL      Unpublished (2003)
COMMENT      Other GSSs: PUPPD96TB
Contact: Cathy Whitelaw
TIGR         9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES
source       1..786
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA0701024"
/clone_lib="ZM 0.6-1.0 KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

```

```

ORIGIN
Query Match          4.0%; Score 67; DB 29; Length 786;
Best Local Similarity 100.0%; Pred. No. 1.8e-20;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1037 AAGAGTGAGAGAGAACTGATCTTTGTTACTATATGCTGCTCTCCAAAGGTTACA 1096
      |||||||
DB 519 AAGAGTGAGAGAGAACTGATCTTTGTTACTATATGCTGCTCTCCAAAGGTTACA 460
      |||||||
QY 1097 TGATATG 1103
      |||||||
DB 459 TGATATG 453
      |||||||

RESULT 10
LOCUS CG062075 858 bp DNA linear GSS 19-AUG-2003
DEFINITION PURPD96TB_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMMBTa0701024,
            genomic survey sequence.
ACCESSION CG062075
VERSION CG062075.1 GI:33934255
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 858)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
            Maize Genomics Consortium
            Unpublished (2003)
            Other GSSs: PUPPD96TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
            Location/Qualifiers
            1..858
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMBTa0701024"
            /clone_lib="ZM_0.6_1.0_KB"
            /notes="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
            Cot selected genomic DNA library"

FEATURES
            source
            1..858
            |||
            22 GCATG 19

ORIGIN
Query Match          4.0%; Score 67; DB 29; Length 858;
Best Local Similarity 100.0%; Pred. No. 1.7e-20;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1037 AAGAGTGAGAGAGAACTGATCTTTGTTACTATATGCTGCTCTCCAAAGGTTACA 1096
      |||||||
DB 444 AAGAGTGAGAGAGAACTGATCTTTGTTACTATATGCTGCTCTCCAAAGGTTACA 503
      |||||||
QY 1097 TGATATG 1103
      |||||||
DB 504 TGATATG 510
      |||||||

RESULT 11
CG193723/c
LOCUS CG193723 438 bp DNA linear GSS 21-AUG-2003
DEFINITION PUJDJ37TB_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMMBTa0642H02,
            genomic survey sequence.

```

```

ACCESSION CG193723
VERSION CG193723.1 GI:34084784
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 438)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
            Maize Genomics Consortium
            Unpublished (2003)
            Other GSSs: PUJDJ37TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
            Location/Qualifiers
            1..438
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMBTa0642H02"
            /clone_lib="ZM_0.6_1.0_KB"
            /note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
            Cot selected genomic DNA library"

ORIGIN
Query Match          3.9%; Score 64; DB 29; Length 438;
Best Local Similarity 100.0%; Pred. No. 6e-19;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1083 CTCCTCAAGGTTACATGATGGGATCTCTCTCTATTATTATAGACAAACTAGGGTTTCA 1142
      |||||||
DB 82 CTCCTCAAGGTTACATGATGGGATCTCTCTCTATTATTATAGACAAACTAGGGTTTCA 23
      |||||||
QY 1143 GCATG 1146
      |||
DB 22 GCATG 19

RESULT 12
CG193725
LOCUS CG193725 446 bp DNA linear GSS 21-AUG-2003
DEFINITION PUJDJ37TD_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMMBTa0642H02,
            genomic survey sequence.
ACCESSION CG193725
VERSION CG193725.1 GI:34084786
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 446)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
            Maize Genomics Consortium
            Unpublished (2003)
            Other GSSs: PUJDJ37TB
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208

```

Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source
1. .446
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBTA0642H02"
/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN
Query Match 3.9%; Score 64; DB 29; Length 446;
Best Local Similarity 100.0%; Pred. No. 5.9e-19; Indels 0; Gaps 0;
Matches 64; Conservative 0; Mismatches 0

Qy 1083 CTCGAAAGGTTACATGATGGGATCTCTCTCTATTATAGACAAACTAGGGTTTCA 1142
|||||
Db 357 CTCGAAAGGTTACATGATGGGATCTCTCTCTATTATAGACAAACTAGGGTTTCA 416
|||||

Qy 1143 GGCA 1146
|||||
Db 417 GGCA 420
|||||

RESULT 13
CC352433/c
LOCUS
DEFINITION
PUDH597D ZM 0.6 1.0 KB Zea mays genomic clone ZMMBTA429I22,
genomic survey sequence.
ACCESSION
CC352433
VERSION
CC352433.1 GI:30821833
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 732)
AUTHORS
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source
1. .732
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBTA429I22"
/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN
Query Match 3.9%; Score 64; DB 28; Length 732;
Best Local Similarity 100.0%; Pred. No. 4.9e-19; Indels 0; Gaps 0;
Matches 64; Conservative 0; Mismatches 0

Qy 1083 CTCGAAAGGTTACATGATGGGATCTCTCTCTATTATAGACAAACTAGGGTTTCA 1142
|||||
Db 357 CTCGAAAGGTTACATGATGGGATCTCTCTCTATTATAGACAAACTAGGGTTTCA 416
|||||

Qy 1143 GGCA 1146
|||||
Db 417 GGCA 420
|||||

RESULT 14
CC006284
LOCUS
DEFINITION
PUDIN46TD ZM 0.6 1.0 KB Zea mays genomic clone ZMMBTA219G20,
genomic survey sequence.
ACCESSION
CC006284
VERSION
CC006284.1 GI:29384846
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 884)
AUTHORS
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source
1. .884
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBTA219G20"
/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN
Query Match 3.9%; Score 64; DB 28; Length 884;
Best Local Similarity 100.0%; Pred. No. 4.6e-19; Indels 0; Gaps 0;
Matches 64; Conservative 0; Mismatches 0

Qy 1130 AACTAGGGTTTCAGGCATATGGCCACATAGGCCTTCTCTGCCCCAAGAGTTTCTTAA 1189
|||||
Db 812 AACTAGGGTTTCAGGCATATGGCCACATAGGCCTTCTCTGCCCCAAGAGTTTCTTAA 871
|||||

Qy 1190 CACT 1193
|||||
Db 872 CACT 875
|||||

RESULT 15
CG096594
LOCUS
DEFINITION
PUFUD537B ZM 0.6 1.0 KB Zea mays genomic clone ZMMBTA0733J10,
genomic survey sequence.
ACCESSION
CG096594
VERSION
CG096594.1 GI:33978888
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 844)

AUTHORS

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J., and
Bennetzen, J.

TITLE

Maize Genomics Consortium

JOURNAL

Unpublished (2003)

COMMENT

Other GSSs: PUFUD53TD

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

FEATURES

source

Location/Qualifiers

1..844

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMETA0733J10"

/clone_lib="ZM 0.6 1.0 KB"

/note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high

Cot selected genomic DNA library"

ORIGIN

Query Match

Best Local Similarity 3.6%; Score 60; DB 29; Length 844;

Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1033

AGACAGAGCTGTAGAGAGGAAGTCTTTGTTACTATATGTTGCTGCTCTCCAAAGGT 1092

Db 476

AGACAGAGCTGTAGAGAGGAAGTCTTTGTTACTATATGTTGCTGCTCTCCAAAGGT 535

QY 1093

TACATGATATGGGATCTCTCTCTATTATAGACAAAAGTGGGTTTCAG 1143

Db 536

TACATGATATGGGATCTCTCTCTATTATAGACAAAAGTGGGTTTCAG 586

Search completed: September 2, 2004, 14:41:27

Job time : 3930 secs